

인구 정신건강 측정

Measuring Population Mental Health



인구 정신건강 측정

Measuring Population Mental Health

(발간) **OECD** (번역) **OECD Korea Policy Centre**

본 문서 및 본 문서에 포함된 모든 데이터와 지도는 특정 영토의 지위 또는 그에 대한 주권, 국가 간 국경 및 경계의 결정, 그리고 특정 영토, 도시 또는 지역의 명칭에 대한 권리를 침해하지 않습니다.

이스라엘에 관한 통계 데이터는 해당 이스라엘 당국의 책임하에 제공됩니다. OECD의 상기 데이터 사용은 국제법에 따른 골란고원, 동예루살렘 및 요르단강 서안지역 이스라엘 정착촌의 지위를 침해하지 않습니다.

OECD 및 유럽연합에 가입한 모든 유럽연합 회원국 참고 사항

사이프러스 공화국은 튀르키예를 제외한 모든 유엔 회원국에서 국가로 인정하고 있습니다. 본 문서의 정보는 사이프러스 공화국 정부가 실효적으로 통제하는 지역에 관한 것입니다.

이 번역본의 원본은 다음과 같은 제목으로 발간되었습니다:

Measuring Population Mental Health © OECD 2023, <https://doi.org/10.1787/5171eef8-en>

이 번역본은 OECD에 의해 만들어진 것이 아니며 OECD의 공식 번역본이 아닙니다.

번역의 질과 원문과의 일관성은 온전히 번역본 저자의 책임 아래 있습니다.

원본과 번역본이 일치하지 않는 경우에는 원본만이 유효한 것으로 간주합니다.

© 2023 OECD/Korea Policy Centre for this translation

서문

정신건강은 현대 사회 및 경제 측면에서 공공의 우선순위로 인식되고 있으며, 그 중요성이 점점 더 강조되고 있다. 정신건강의 긍정적인 측면과 부정적인 측면을 이해하고 매핑하는 것은 민간 부문과 시민 사회에서 시행 중인 다양한 공공 정책과 조치에 대해 파악하는 데 핵심적인 역할을 한다. 본 보고서는 국가 통계청과 기타 데이터 생산기관에서 조사 빈도를 높이고 국제적으로 통일되고 일관된 방식으로 조사를 수행하여 고품질의 인구 정신건강 결과를 취합할 수 있도록 지원하는 것을 목표로 한다. 본 보고서는 OECD 국가의 기존 측정 관행을 문서화하고 기존 측정 도구의 장점과 한계점을 논의하며 가정 및 사회의 정신건강에 관한 설문조사에서 채택할 측정의 우선순위(부정적 정신건강 및 긍정적 정신건강에 대한)에 대해 권장사항을 제공한다. 본 보고서에서는 완전한 측정 지침을 제시하지는 않지만 대중의 웰빙에 대한 통계 의제를 선도하는 과거 OECD WISE 센터의 업무를 보완하는 역할을 하며, 주관적 웰빙, 신뢰, 근무 환경의 질에 대한 지침을 포함한다. 본 간행물은 OECD How's Life? 연속 간행물에 포함되어 있는 정신건강 및 웰빙에 대한 특별 평가의 일환으로 작성된 두 개의 보고서 중 첫 번째 보고서이다. 두 번째 보고서는 OECD 웰빙 프레임워크에 예시된 바와 같이 대중의 삶에 내재된 다양한 경제적, 사회적, 환경적 및 관계적 관점과의 상호연관성을 조사하여 인구 정신건강을 웰빙의 관점에서 다각도로 조명한다.

본 보고서는 OECD WISE 센터에서 작성하였다. 본 보고서는 Lara Fleischer(제1장, 제2장)와 Jessica Mahoney(제2장, 제3장) 저자로 구성되었으며, 통계 작업은 Jessica Mahoney의 주도하에 진행되었다. 본 보고서에 제공된 정보에 대한 선행 연구를 수행한 Manuela Grabosch, Muriel Levy, Nikita Arora에게 감사를 표한다. 본 프로젝트의 콘텐츠 편집은 Lara Fleischer의 주도로 진행되었으며, Carrie Exton이 감독하고 Marco Mira d'Ercole과 Romina Boarini가 추가로 편집 작업을 수행했고, Romina Boarini의 지도하에 출판되었다. Martine Zaida는 커뮤니케이션 담당자로 프로젝트 전반에 걸쳐 필수적 지원을 제공했다. Cassandra Morley는 출판을 위한 원고를 준비하고 서식을 구성했으며, Patrick Hamm이 원고 교정 및 교열 작업을 수행하고, Sonia Primot가 앞표지를 디자인했다.

본 보고서는 OECD 통계정책위원회(Committee on Statistics and Statistical Policy)와 OECD 보건위원회(Health Committee)의 각국 대표가 제공한 유용한 의견에서 많은 도움을 받았다. 그들의 기여와 조언에 깊은 감사를 표한다.

또한 초안 작성이나 특정 문의 사항에 대해 데이터를 제공하고 다양한 의견과 통찰력을 제시해 준 내외부 참여자들에게도 감사 인사를 전한다. 여기에 모두 언급할 수는 없지만 본 보고서의 발간을 위해 도움을 준 모든 분들에게 감사를 표한다.

- Doron Wijker, Christopher Prinz, Shunta Takino, Emily Hewlett(OECD 고용노동사회국)
- 본 프로젝트를 지원하는 비공식 자문그룹의 자문위원:
 - Adam Coutts(하버드대학교 웨더헤드 국제관계센터 선임연구위원, 케임브리지대학교 모듈린 컬리지)
 - Alexandra Lazaro, Cath Davies, Jess Rackham, Lily Makurah, Linda Bullivant(영국 보건사회복지부)

- Amelia Walters(빅토리안 정신건강 및 웰빙 협력 센터 이사)
- Andrew Steptoe(유니버시티칼리지런던 심리학 및 역학 교수, 행동 과학 및 보건 학과장)
- Angeline Ferdinand(멜버른대학교 멜버른 인구 및 세계보건학교 보건정책센터 연구원)
- Bernard Jacob(벨기에 연방공공서비스 - 공중보건 벨기에 정신건강 관리개혁 연방조정관)
- Caroline Cohrdes(로버트코흐연구소 정신건강부)
- Catherine Carty(UNESCO 석좌 프로젝트 매니저)
- Claire Gibbons(로버트우드존슨재단 수석 프로그램 담당관)
- Daniel Hugh Chisholm(세계보건기구 유럽지역사무소 정신건강 프로그램 관리자)
- David Finch(보건재단 Health Lives팀 선임 담당관)
- Elisabeth Ng Langdal(정신건강 및 인권정보 책임자)
- Emma Lawrance, Jessica Newberry Le Vay(임페리얼칼리지런던 글로벌보건혁신연구소)
- Fredrik Lindencrona(스웨덴 지방자치단체협회 전략개선팀 책임자)
- George Ploubidis(유니버시티칼리지런던 사회연구소 인구보건 및 통계학 교수)
- Gregory Bratman(워싱턴대학교 환경 및 웰빙 연구소장, 조교수)
- Hannes Kröger(전 독일경제연구소 SocPsych-MH 부서장)
- Harold Alan Pincus(컬럼비아대학교 바젤로스 의과대학 어빙 연구소 공동 연구소장, 정신의학과 교수 겸 연구 부의장)
- Hedinn Unnsteinsson(아이슬란드 총리실 수석정책분석관, 아이슬란드 정신건강연합 회장)
- Katie Hayes(캐나다 보건부 수석정책분석관)
- Layla McCay(도시 디자인 및 정신건강센터 센터장)
- Lene Søvdal(임상 심리학자, 정신건강 고문 및 연구원)
- Luca Bernardi(리버풀대학교 정치학 전임 강사)
- Margaret Walker, Peter Falkai(유럽정신의학협회 집행이사 겸 회장)
- Patricio V Márquez(존스홉킨스대학교 블룸버그 공중보건대학 선임연구원, 전 세계은행그룹 공중보건 수석연구원)
- Richard Layard(런던 정치경제대학교 경제성과센터 공동체 웰빙 프로그램 공동 책임자)
- Ronni M. Greenwood(리머릭대학교 심리학과 강사 및 사회공동체 심리학자)
- Sarah Hinde(호주 보건부 정신건강 데이터 및 자료 태스크포스 차관보)
- Shekhar Saxena(하버드 공중보건대 글로벌 정신건강 실무 교수)

- Tyler Norris(웰빙 트러스트 CEO, MDiv)
- Ziggi Ivan Santini(덴마크 국립보건원 연구원)
- 2022년 4월 28일 국가 통계청을 대상으로 개최된 프로젝트 웨비나 참가자:
 - Adriana Pérez Amador, Olinca Páez(멕시코 통계청)
 - Alessandra Tinto, Lidia Gargiulo(이탈리아 국가 통계청)
 - Camélia St-Denis, Emma Howieson, Isabelle Lévesque, Joie Huggins, Leah Smith(캐나다 통계청)
 - Chika Arita(일본 총무성)
 - Claudio Peter, Daniela Schuler(스위스 보건감시기구)
 - Emilie Hegelund(덴마크 통계청)
 - Kali Kong(미국 경제분석국)
 - Kelly Sabbe(벨기에 통계청)
 - Dace Krievkalne(라트비아 통계청)
 - Guy Weber(룩셈부르크 보건국)
 - Lydia Gisle(벨기에 보건연구소)
 - Matthew Montgomery(호주 통계청)
 - Mónika Földvári(헝가리 통계청)
 - Naama Rotem(이스라엘 중앙 통계국)
 - Sara Holsbrink(스웨덴 공중보건국)
 - Rosario Gonzalez Garcia(스페인 통계청)
 - Sunhye Choi(대한민국 질병관리청)

본 보고서는 로버트우드존슨재단의 지원을 받아 작성되었다. 본 문서에 언급된 의견과 논의가 반드시 재단의 견해를 반영하는 것은 아니다.

목차

| | |
|---|-----------|
| 서문 | 3 |
| 독자를 위한 가이드 | 11 |
| 요약 | 13 |
| 1. 인구 정신건강이란 무엇이며 왜 측정해야 하는가? | 15 |
| 인구 정신건강 결과에 중점을 두는 것의 중요성 | 16 |
| 정신건강의 개념: 질병에서 건강까지 | 18 |
| 정신건강 문제 | 23 |
| 긍정적 정신건강 | 26 |
| 결론 | 27 |
| 참고문헌 | 28 |
| 주 | 33 |
| 2. 인구 정신건강 측정: 측정 도구 및 기존 국가별 측정 관행 | 35 |
| 인구 정신건강을 측정하는 데 사용할 수 있는 도구는 무엇인가? | 37 |
| OECD 국가에서 기존에 수집하고 있는 인구 정신건강 데이터는 무엇인가? | 45 |
| 결론 및 향후 방향 | 54 |
| 부속서 2.A. 국가별 정신건강 결과 측정 조사 | 55 |
| 부속서 2.B. 정신건강을 측정하기 위한 표준화된 조사 도구에 대한 세부 정보 | 71 |
| 참고문헌 | 82 |
| 주 | 90 |
| 3. 가구 조사를 통한 인구 정신건강 측정의 모범 사례 | 91 |
| 통계 품질 | 95 |
| 데이터 수집 | 116 |
| 분석 | 121 |
| 결론 | 125 |
| 참고문헌 | 128 |
| 주 | 141 |

그림차례

| | |
|--|-----|
| 그림 1.1. OECD 웰빙 프레임워크 | 17 |
| 그림 1.2. 정신건강의 모델 | 19 |
| 그림 1.3. 강과 같은 정신건강 연속체 | 20 |
| 그림 1.4. 호주 아동의 정신건강의 이중 연속체 | 22 |
| 그림 2.1. 정신건강 측정 도구별 응답 부담과 정확도 간의 관계 | 41 |
| 그림 2.2. 일반적으로 선별 도구는 구조화된 면담에 비해 주요우울장애 유병율이 높게 나타난다 | 43 |
| 그림 2.3. 선별 도구에 의해 심리적 고통을 경험한 적이 있다고 확인된 응답자의 비율은 심리적 고통을 경험한 적이 있다고 확인된 응답자의 비율보다 훨씬 낮은 것으로 나타난다 | 44 |
| 그림 2.4. 대부분의 OECD 국가는 건강 조사 및 일반 사회 조사를 통해 인구 정신건강을 측정한다고 보고한다 | 46 |
| 그림 2.5. 많은 OECD 국가에서는 정신건강 데이터를 수집하는 빈도가 낮으며 절반 이상이 4년~ 10년에 한 번 주기로 설문조사를 실시하는 것으로 보고했다 | 47 |
| 그림 2.6. 모든 OECD 국가에서 정신건강 문제 및 긍정적 정신건강에 대한 데이터를 수집한다고 응답했으며 후자는 주로 생애 평가에 중점을 두고 있다 | 49 |
| 그림 2.7. OECD 국가는 선별 도구와 증상 및 장애 경험에 관한 질문이 정신건강 문제를 측정하는데 가장 일반적으로 사용하는 도구라고 보고했다 | 50 |
| 그림 2.8. 일반적인 심리적 고통과 우울증 증상을 포착하는 선별 도구가 불안이나 기타 장애의 증상을 포착하는 선별도구에 비해 더 일반적으로 사용된다 | 51 |
| 그림 2.9. 국가에서는 불안 데이터를 수집하지만 표준화되지 않은 측정 방법을 사용하는 경우가 많다 | 52 |
| 그림 2.10. 정서 데이터는 행복 데이터보다 수집되는 빈도가 높지만, OECD 국가에서 삶의 만족도를 넘어 긍정적 정신건강에 대한 데이터를 수집하기 위해 사용하는 도구가 통일되어 있지 않다 | 53 |
| 그림 3.1. 정신건강 관련 문항의 경우 다른 문항에 비해 무응답률이 높으며, 이는 국가별로 큰 차이를 보인다 | 104 |
| 그림 3.2. 청년층, 남성 및 교육 수준이 낮은 집단은 정신건강 관련 문항에 응답할 가능성이 낮다 | 106 |
| 그림 3.3. 심리적 고통 및 우울 증상 위험 유병률의 경우 OECD 국가 중 유럽 국가 간에 100%까지 차이가 난다 | 108 |
| 그림 3.4. 성별과 인종/민족별로 보고된 결과의 차이는 근본적인 정신건강의 차이 때문인가 아니면 측정 문제에서 비롯되었나? | 109 |
| 그림 3.5. 팬데믹 첫해에 OECD 국가 중 유럽 8개국에서 우울증 증상이 크게 증가했다 | 113 |
| 그림 3.6. 2019년까지 OECD 유럽 국가에서는 정신건강이 다소 개선되었으며 미국은 정신건강에 대한 인식이 개선되었음에도 대체로 안정적인 수준을 유지했다 | 115 |
| 그림 3.7. 대면 조사에서 전화 조사로의 전환은 부정적 정서 균형의 악화와 관련된다 | 119 |
| 그림 3.8. 긍정적 정신건강 척도는 정규 분포에 더 가까울 수 있다 | 122 |
| 그림 3.9. 채점 방식이 다르면 유병률 추정치가 달라질 수 있다 | 124 |

표목차

| | |
|---|----|
| 표 1. 국가 및 지역별 ISO 코드 | 11 |
| 표 2. 보고서에 사용된 약어 목록 | 11 |
| 표 1.1. 긍정적 정신건강 달성 정도에 영향을 미치는 정신건강 상태의 차이 | 22 |
| 표 2.1. 정신건강 측정 도구별 장점 및 한계점 | 41 |
| 표 2.2. 여러 국가에서 아동 및 청소년의 정신건강에 대한 설문조사 또는 설문조사 모듈을 도입했다 | 48 |

| | |
|---|-----|
| 표 2.3. OECD 국가의 절반 이상이 코로나19 팬데믹 기간에 정신건강 데이터의 수집 빈도를 증가시켰다고 보고했다 | 48 |
| 표 2.4. 정신건강 상태를 모니터링하기 위한 구조화된 면담 개요 | 55 |
| 표 2.5. 일반적인 정신 질환과 특정 정신건강 상태의 위험을 모니터링하기 위한 타당화된 선별도구 개요 | 55 |
| 표 2.6. 이전 진단에 관한 문항 개요 | 58 |
| 표 2.7. 증상 및 정신건강 상태의 경험에 관한 문항 개요 | 60 |
| 표 2.8. 자살 사고 및 자살 시도에 관한 문항 개요 | 63 |
| 표 2.9. 일반적인 정신건강 상태에 관한 문항 개요 | 64 |
| 표 2.10. 긍정적 정신건강 지표 개요 | 66 |
| 표 2.11. 국가별로 수집된 기타 정신건강 관련 주제에 대한 개괄적인 개요 | 69 |
| 표 2.12. MHI-5 설문지(점수표 포함) | 72 |
| 표 2.13. 케슬러 척도 10/6 설문지(점수표 포함) | 73 |
| 표 2.14. PHQ-9/8 설문지(점수표 포함) | 74 |
| 표 2.15. GAD-7/GAD-2 설문지(점수표 포함) | 75 |
| 표 2.16. PHQ-4 설문지(점수표 포함) | 75 |
| 표 2.17. WG-SS 강화 설문지 | 76 |
| 표 2.18. AUDIT/AUDIT-C 설문지(점수표 포함) | 77 |
| 표 2.19. 주관적 웰빙에 관한 OECD 핵심 문항 | 78 |
| 표 2.20. WHO-5 설문지(점수표 포함) | 78 |
| 표 2.21. SF-36 활력 하위 척도 설문지(점수표 포함) | 79 |
| 표 2.22. SWLS 설문지(점수표 포함) | 79 |
| 표 2.23. MHC-SF 설문지(점수표 포함) | 80 |
| 표 2.24. (S)WEMWBS 설문지(점수표 포함) | 81 |
| 표 3.1. 통계 품질, 데이터 수집 및 분석 지표에 대한 정신건강 도구 성능 개요 | 93 |
| 표 3.2. 낙인과 유병률 간의 관계는 복잡하지만 경우에 따라 낙인으로 인해 정신장애 유병률이 낮은 수준으로 보고될 수 있다 | 105 |
| 표 3.3. 유사한 문항에 대한 응답의 상관관계는 문구가 긍정적인지 또는 부정적인지에 따라 달라진다 | 117 |

독자를 위한 가이드

표 1. 국가 및 지역별 ISO 코드

| | | | | | |
|-----|-------|-----|-------|------|---------|
| AUS | 호주 | FIN | 핀란드 | LVA | 라트비아 |
| AUT | 오스트리아 | FRA | 프랑스 | MEX | 멕시코 |
| BEL | 벨기에 | GBR | 영국 | NLD | 네덜란드 |
| CAN | 캐나다 | GRC | 그리스 | NOR | 노르웨이 |
| CHE | 스위스 | HUN | 헝가리 | NZL | 뉴질랜드 |
| CHL | 칠레 | IRL | 아일랜드 | OECD | OECD 평균 |
| COL | 콜롬비아 | ISL | 아이슬란드 | POL | 폴란드 |
| CRI | 코스타리카 | ISR | 이스라엘 | PRT | 포르투갈 |
| CZE | 체코 | ITA | 이탈리아 | SVK | 슬로바키아 |
| DEU | 독일 | JPN | 일본 | SVN | 슬로베니아 |
| DNK | 덴마크 | KOR | 한국 | SWE | 스웨덴 |
| ESP | 스페인 | LTU | 리투아니아 | TUR | 튀르키예 |
| EST | 에스토니아 | LUX | 룩셈부르크 | USA | 미국 |

표 2. 보고서에 사용된 약어 목록

| | |
|----------|---|
| ACASI | 오디오와 컴퓨터를 보조적으로 활용한 자기면접(Audio computer-assisted self-interviewing) |
| ADHD | 주의력결핍 과잉행동장애(Attention deficit hyperactivity disorder) |
| ASD | 자폐 스펙트럼 장애(Autism spectrum disorder) |
| AUDIT | 알코올사용장애 식별 테스트(Alcohol Use Disorders Identification Test) |
| AUDIT-C | 알코올사용장애 식별 테스트 - 단축형(Alcohol Use Disorders Identification Test - Concise) |
| CAGE | CAGE 약물남용선별도구(CAGE Substance Abuse Screening Tool) |
| CES-D | 역학연구 우울 척도(Epidemiological Studies Depression Scale) |
| CFA | 확인적 요인분석(Confirmatory factor analysis) |
| CIDI | 복합 국제 진단 면담(Composite International Diagnostic Interview) |
| CIUS | 강박적 인터넷 사용 척도(Compulsive Internet Use Scale) |
| COVID-19 | 코로나바이러스감염증-19(Coronavirus disease of 2019) |
| DASS | 우울, 불안 및 스트레스 척도(Depression, Anxiety and Stress Scale) |
| DAWBA | 발달 및 웰빙 평가 - 섭식장애 모듈(Development and Wellbeing Assessment - Eating Disorder Module) |
| DAWMA | 발달 및 웰빙 평가(Development and Well-being Assessment) |
| DSM | 정신질환 진단 및 통계 편람(Diagnostic and Statistical Manual of Mental Disorders) |
| EPDS | 에딘버러 산후우울증 척도(Edinburgh Post-natal Depression Scale) |
| EFA | 탐색적 요인분석(Exploratory factor analysis) |
| EHIS | 유럽건강면접조사(European Health Interview Survey) |
| EU-SILC | 유럽 소득 및 생활수준 통계(European Union Statistics on Income and Living Conditions Survey) |
| EVI | 에너지 및 활력 지수(Energy and Vitality Index) |
| GAD | 범불안장애(Generalized Anxiety Disorder) |

| | |
|-------------|--|
| GHQ | 일반 건강 설문지(General Health Questionnaire) |
| HADS | 홉킨스 증상 체크리스트(Hopkins Symptom Checklist) |
| HRQoL | 건강관련 삶의 질(Health-related quality of life) |
| IES-R | 사건 충격 척도 - 개정판(Impact of Event Scale - Revised) |
| I-PAPI | 면담자 시행 지필 설문조사(Interviewer-administered paper and pencil surveys) |
| K10 (K6) | 케슬러 척도(Kessler Scale) |
| MDE | 주요우울장애(Major depressive disorder) |
| MHC | 정신건강 요인 요약(Mental health component summary) |
| MHC-LF | 정신건강 연속성 척도 - 확장형(Mental Health Continuum - Long Form) |
| MHC-SF | 정신건강 연속성 척도 - 단축형(Mental Health Continuum - Short Form) |
| MHI | 정신건강검사(Mental Health Inventory) |
| Mini-SPIN | 축약형 사회공포증검사(Mini-Social Phobia Inventory) |
| NHS | 국립보건원(National Health Service) |
| NSO | 통계청(National statistical office) |
| OECD | 경제협력개발기구(Organisation for Economic Co-operation and Development) |
| PCL | 외상후 스트레스 장애 체크리스트(PTSD Checklist) |
| PC-PTSD-5 | DSM-5 기준 외상 후 스트레스 장애 선별 검사(Primary Care PTSD Screen for DSM-5) |
| PCS | 신체적 요소 요약(Physical component summary) |
| PHQ | 환자 건강 설문지(Patient Health Questionnaire) |
| PHQ-ED | 환자 건강 설문지 - 섭식장애 모듈(Patient Health Questionnaire - Eating Disorder Module) |
| PHQ-PD | 환자 건강 설문지 - 공황 장애(Patient Health Questionnaire - Panic Disorder) |
| PSS | 코헨 지각된 스트레스 척도(Cohen Perceived Stress Scale) |
| PTSD | 외상 후 스트레스 장애(Post-traumatic stress disorder) |
| SCARED | 아동 불안 및 관련 정서 장애 선별 척도(Screen for Child Anxiety and Related Emotional Disorders) |
| SCOFF | 섭식장애 설문지(eating disorder questionnaire) - 구토 경험, 식이 조절 실패에 대한 두려움, 급격한 체중 감량, 똥똥하다는 인식, 음식에 대한 집착(Sick, Control, One, Fat, Food) |
| SDQ | 강점·어려움 설문지(Strengths and Difficulties Questionnaire) |
| SF-12 (-36) | 단축형 건강 상태(Short Form Health Status) |
| SMFQ | 단축형 기분 및 감정 설문지(Short Moods and Feelings Questionnaire) |
| SRMH | 자기보고 정신건강(Self-reported mental health) |
| STAI | 상태-특성 불안 척도(The State and Trait Anxiety Scale) |
| (S)WEMWBS | (단축형) 워릭-에딘버러 정신적 웰빙 척도 지수[(Short) Warwick-Edinburgh Mental Well-being Scale] |
| SWLS | 삶의 만족도 척도(Satisfaction with Life Scale) |
| ROC | ROC 곡선(Receiver operating characteristic curve) |
| WG-ES | 기능에 관한 워싱턴 그룹 확장형 설문지(Washington Group Extended Set on Functioning) |
| WG-SS | 장애통계에 관한 워싱턴 그룹 간이형 설문지(Washington Group on Disability Statistics Short Set on Functioning) |
| WHO | 세계보건기구(World Health Organization) |
| WHO-5 | WHO-5 웰빙지수(WHO-5 Wellbeing Index) |
| WHOQOL-BREF | WHO 삶의 질 - 간편형 척도(WHO Quality of Life - BREF) |

요약

정신건강은 사람들의 웰빙 개선에 필수적인 요소이다...

정신건강은 사람들의 삶에서 중추적인 역할을 하며 본질적으로 광범위한 웰빙의 다른 많은 측면과 관련이 있다. 특히 코로나19 팬데믹 기간에 정신건강의 중요성이 강조되었는데, 이는 코로나19가 공중보건에 직접적으로 악영향을 미치고 인명 손실을 초래함과 동시에 사회적 고립, 실직 및 재정적 불안정이 심화되면서 사람들의 정신건강이 크게 악화되었기 때문이다. OECD 15개국의 데이터에 따르면 2020년 말까지 인구의 1/4 이상이 우울이나 불안 증상을 경험한 것으로 나타났다. 이미 팬데믹이 도래하기 훨씬 전부터 인구의 절반이 일생에서 한 번 이상 정신건강 문제를 경험할 것으로 예측되었으며, 정신건강 문제로 인한 경제적 비용은 연간 GDP의 4% 이상에 달할 것으로 추정되었다. 반면 긍정적 정신건강은 스트레스에 대한 회복탄력성을 향상시키고 사람들이 자신의 목표를 깨닫고 공동체에 적극적으로 기여하도록 도울 수 있다. 따라서 OECD 각국의 보건부 및 기타 정부기관에서는 긍정적 정신건강, 즉 높은 수준의 정서적, 심리적 웰빙을 그 자체로 정책 목표로 인식하는 경우가 점점 더 확대되고 있다.

...그러나 인구 수준에서 정신건강을 모니터링하는 적합한 방법에 대한 지침이 결여되어 있다.

정신건강을 개선하려면 보다 광범위한 인구 수준에서 정신건강 문제와 긍정적 정신건강에 대한 결과를 모니터링하는 것이 필수적이다. 의료 전문가에게 진단 또는 치료를 받은 환자만을 대상으로 작성된 통계는 해당 국가의 의료 시스템에 대한 접근성 및 발전 정도에 따라 크게 영향을 받으며 위험군을 조기에 파악하려면 의료 서비스를 이용하기 전에 정신건강 결과를 추적해야 한다. 또한 긍정적 정신건강은 인구의 기초 자산이므로 그 자체로 추적할 가치가 있다. 정신건강 증진을 위한 성공적인 전략을 수립하려면 사람들의 물질적 조건, 삶의 질, 사회적 관계(및 이러한 요인의 불평등)와 같은 광범위한 위험 및 회복탄력성 요인이 정신건강에 미치는 영향을 이해해야 한다. 해당 주제에 대한 데이터는 일반적으로 정신건강 관련 문항이 포함된 확장형(사회적) 인구 통계 조사를 통해 수집되며, 이러한 데이터는 해당 주제에 대한 이해를 증진하고 정책 개발을 위한 타당한 근거를 제공할 수 있다.

국제적으로 인구 전반의 정신건강 결과에 대한 데이터를 제공하는 경우가 확대되고 있지만, 조사의 빈도가 낮고 국가 간의 통일성이 결여된 실정이다. OECD는 정신건강 개선에 대한 사회 전반의 대응을 촉진하기 위한 노력의 일환으로 인구 정신건강 통계를 정기적으로 발표하고 있으나, 해당 통계 중 일부는 OECD 국가의 일부 국가에 한해 정기적으로 제공되는 데 그치고 있다. 일부 국가의 경우 통계가 작성된 지 5년 이상이 지났으며 경우에 따라 비공식적인 출처를 기반으로 작성된 데이터도 있는 등 여러 한계점을 가진다.

본 보고서의 목적은 무엇인가?

본 보고서는 국가 통계청과 기타 데이터 생산기관에서 조사의 빈도를 높이고 국제적으로 일관성 및 통일성 있는 방식으로 조사를 수행하여, 고품질의 인구 정신건강 결과 측정을 수집할 수 있도록 지원하는

것을 목표로 한다. OECD는 2022년 2월 OECD 통계정책위원회의 설문지를 통해 정신건강과 관련하여 회원국에서 현재 시행하고 있는 일에 대해 조사했다. 거의 모든 국가 통계청에서 보고했으며, 많은 경우 보건기관에서도 답변을 보내왔다. 본 보고서는 각국이 인구 정신건강 결과를 수집하는데 있어 수렴하는 부분과 격차가 존재하는 부분을 파악하기 위해 기존의 측정 관행을 문서화 했다. 또한 이용 가능한 측정 도구를 평가하여 공식 데이터 생산기관이 가정 및 사회의 정신건강에 관한 조사에서 우선적으로 채택할 수 있는 척도에 대한 권장사항을 제공한다.

주요 메시지 및 권장사항

- **인구 수준에서 정신건강 문제와 긍정적 정신건강에 대한 데이터를 수집하면** 정신건강에 대해 보다 완전한 이해를 도모할 수 있다. 인구 조사에 사람들의 삶의 다른 측면에 대한 정보와 관련된 문항을 통합하면 정신건강 결과의 개선을 유도하는 동인 및 정책 수단을 보다 정확히 파악하는 데 도움이 될 것이다. 이는 사후대응적 정책이 아닌 사전예방적 정책을 설계하고, 정신건강 문제를 줄이는 동시에 정서적, 심리적 웰빙을 촉진하는 정신건강 전략을 수립하는 데 새로운 길을 제시할 수 있다.
- **코로나19 팬데믹으로 인해 정신건강 데이터 수집에 대한 새로운 노력이 촉발되었으며, 이러한 모멘텀을 계속 유지해 나가는 것이 중요하다.** 2020년 3월 이후 OECD 10개국 중 7개국에서 기존 설문조사에 정신건강 모듈을 추가하거나 정신건강 설문조사를 새롭게 시작했으며, 이들 중 상당수는 1년에 여러 차례 실시된 것으로 나타났다. 그러나 이러한 조사가 앞으로도 계속될지 여부는 불확실하다. 팬데믹 이전의 일상으로 돌아간다는 것은 OECD 국가의 절반이 4년~10년에 한 번 주기로 정신건강 데이터를 수집한다는 의미이다.
- **정신건강의 일부 측면이 다른 측면에 비해 더 빈번하게 포착되는 경향이 있어 국가 간 측정 통일성을 개선할 여지가 있다.** 특히 불안, 정서 및 행복(eudaimonia)(즉 삶의 의미와 목적의식) 등의 증상을 측정하는 데 있어 통일성이 결여되어 있고, 불안 및 우울 장애를 넘어 특정 정신건강 상태를 평가하는 도구의 사용이 균일하지 않다는 점이 이를 반영한다.
- **본 보고서는 관련 인구 조사에 4가지 특정 도구를 추가하여 국제적으로 통일성 있는 인구 정신건강 지표를 구축할 것을 제안한다.** 이 권장사항은 통계 품질, 응답 부담, 비용, 기존 데이터 수집 관행에 대한 비교 평가를 기반으로 작성되었다. 이는 OECD 국가에서 인구 정신건강 결과를 파악하기 위해 기존에 사용하고 있는 도구의 단계적 폐지를 의미하지는 않는다.
 - **정신건강 문제 - 우선순위 권장사항:** 환자 건강 설문지-4(PHQ-4)는 정신건강 설문조사에서 정기적으로 수집하는 PHQ-8 또는 PHQ-9와 함께 실시 빈도가 높은 설문조사에 포함될 수 있다. 이 설문지는 단 4개의 문항으로 구성되어 있으며, 우울증과 불안 증상을 포괄한다.
 - **긍정적 정신건강 - 권장사항:** 국가별 측정 관행의 경향에 따라 WHO-5 또는 SWEMWBS를 사용하여 긍정적 정신건강의 정서적 측면과 행복 측면을 비교 측정할 수 있다. 정서 및 행복을 구체적으로 측정하는 주제는 향후 주관적 웰빙에 관한 OECD 연구에서 계속 탐구될 것이다.
 - **일반적인 정신건강 상태 - 권장사항:** 일반적으로 응답자의 신체건강을 평가할 때 질의하는 문항과 유사하게 응답자의 일반 정신건강 상태에 대한 단일 문항이 각국의 광범위한 데이터 인프라 시스템 전반에 걸쳐 다양한 설문조사에 포함될 수 있다. OECD 국가 중 절반 이상이 이미 해당 문항을 설문조사에 포함하고 있지만 문항의 문구는 매우 다양하다. OECD 국가 중 최소 3개국에서 다음과 같은 문항을 채택하고 있다. ‘일반적으로 귀하의 정신건강은 어떻습니까? 훌륭하다/매우 좋다/좋다/보통이다/나쁘다.’

1 인구 정신건강이란 무엇이며 왜 측정해야 하는가?

정신건강은 사람들의 웰빙에 필수적인 요소이며, 이를 측정하는 것은 궁극적으로 사람들에게 중요한 요소가 무엇인지 모니터링하는 데 필수적이다. 본 보고서의 목적은 공식 데이터 생산기관에서 인구 정신건강에 대한 데이터 조사를 수행하는 빈도를 높이고 국제적으로 통일성 있는 방식으로 데이터를 수집하도록 장려하는 것이다. 정신건강 문제(진단 결과에 관계없이)에서 긍정적 정신건강 상태에 이르기 까지 정신건강의 모든 측면을 고려하면 정신건강 시스템 및 서비스를 사후대응적 정책이 아닌 사전예방적 정책 설계를 위한 새로운 방법으로 제시할 수 있다. 이는 정신건강 문제를 줄이는 동시에 사람들의 웰빙을 촉진하는 데 초점을 맞춘 정책을 마련할 여지를 제공할 수 있다. 가정 및 사회, 건강관련 설문조사에서 정신건강 문제 및 긍정적 정신건강에 대한 데이터를 수집하면 정신건강에 대한 보다 완전한 이해를 도모할 수 있고, 정신건강 개선을 위한 동인 및 정책 수단을 정확히 파악하는 데 도움이 될 것이다.

긍정적 정신건강은 사람들의 웰빙에 필수적인 요소이다. 긍정적 정신건강은 개인이 자신의 잠재력을 발휘하고 삶의 스트레스에 대처하며 생산적으로 일하고 자신의 공동체에 긍정적인 기여를 할 수 있도록 장려한다(세계보건기구, 2019^[1]). 반면 정신건강 문제는 전 세계적으로 질병 부담이 가장 크고 빠르게 증가하는 범주 중 하나이다. 정신건강 문제로 인한 고용 및 생산성 저하와 정신건강 시스템에 대한 투자를 포함한 경제적 비용은 OECD 국가 GDP의 4% 이상에 달하는 것으로 추정된다(Rehm 및 Shield, 2019^[2], OECD, 2021^[3]). 코로나19 팬데믹이 도래하기 훨씬 전인 2016년 (자살, 급성 알코올 남용 또는 약물 과다 복용으로 인한) 절망사(deaths of despair)는 OECD 국가에서 예방 가능성이 가장 높은 사망 원인 중 하나로, 살인으로 인한 사망보다 평균 6배, 교통사고로 인한 사망보다 3배 더 높았다(OECD, 2020^[4]). 2016년~2018년의 기간에 걸쳐 OECD 국가에서 거주하는 8명 중 1명은 일상적인 하루 동안 긍정적인 감정보다 부정적인 감정을 더 많이 경험하는 것으로 나타났다(OECD, 2020^[4]).

특히 코로나19 동안 정신건강이 가장 중요하게 공론화되었다. 코로나19로 인한 직접적인 영향 외에도 사회적 고립, 실직, 재정적 불안정이 심화되면서 사람들의 정신건강이 심각하게 악화되었고, 이는 높은 사망자 수에 영향을 미친 것으로 나타났다. 실제로 2020년 말까지 OECD 15개국에서 인구의 4분의1 이상이 불안 또는 우울 증상을 경험한 것으로 확인되었다(OECD, 2021^[3], OECD, 2021^[5]). 특히 여성, 청년, 불안정한 고용 및 재정 상황에 처한 사람, 소수 민족 또는 인종, 기존에 정신건강에 문제가 있거나 약물 남용 장애를 가진 사람을 포함한 취약계층이 영향을 받았다.

정신건강은 사람들의 웰빙에 중요한 요소이며 인구의 상당수가 정신건강 문제를 겪고 있거나 부정적 정신건강의 영향을 받고 있다는 점은 분명하다. 하지만 지금까지 정부가 보다 광범위한 인구 수준에서 정신건강을 모니터링하는 적합한 방법과 정신건강 문제 및 긍정적 정신건강을 모두 고려하는 방법에 대한 논의가 충분히 이루어지지 않았다. 이를 위해서는 ‘정신건강’의 정확한 의미를 파악하고 질병 치료, 예방 및 건강 증진 전략을 담당하는 정책 입안자들과 정책 관련성이 가장 높은 결과에 대한 논의가 필요하다.

본 장에서는 먼저 인구 수준의 정신건강 결과 측정을 정기적으로 수행하고 데이터를 수집해야 하는 이유에 대해 설명한다. 그 다음 정신건강 문제와 긍정적 정신건강을 포함하여 정신건강의 여러 구성요소를 연구와 실무에서 식별하는 방법에 대해 설명한다¹. 다음 장에서는 OECD 국가에서 사용 가능한 측정 도구 및 현재 국가별 측정 활동(제2장)과 기존 통계 품질 및 측정 관행(제3장)에 대해 다룬다. 본 보고서는 보고서 이용자들의 공통적인 이해를 돕기 위해 보고서에 사용된 용어 및 관련 내용에 대한 설명을 제공한다.

인구 정신건강 결과에 중점을 두는 것의 중요성

OECD는 오랜 기간 국제 보건 통계를 수집하고 정신건강 개선을 위한 사회 전반의 대응을 촉구해 왔다. 이러한 노력의 일환으로 2015 *OECD Recommendation on Integrated Mental Health, Skills and Work Policy* (2015년 OECD 통합 정신건강, 기술 및 일자리 정책에 관한 권고안)과 그 후속 보고서인 *Fitter Minds, Fitter Jobs*(더 건강한 마음, 더 좋은 일자리)를 발간했다. 최근에는 정신건강 성과를 파악하기 위한 프레임워크를 제시하고 각국이 보건 시스템 성과에 관한 정책 및 서비스를 제공하고 있는지를 평가하는 *A New Benchmark for Mental Health Systems*(새로운 정신건강 시스템 벤치마크)를 발간했다(OECD, 2021^[3], OECD, 2015^[6], OECD, 2021^[7]). 정신질환 예방, 정신적 웰빙 증진, 정신건강에 대한 다각적인 접근은 OECD의 새로운 벤치마크 프레임워크의 핵심 원칙 중 하나이다. 이러한 원칙에 따라 삶의 만족도, 자살률, 교육 및 고용 상태에 따른 정신적 고통 불평등 등 여러 인구 수준의 결과 지표가 포함된다. 또한 OECD *How's Life?* 보고서(40개 이상의 회원국 및 파트너국의 웰빙, 불평등 및 지속가능성을 평가하는 보고서, 상자 1.1 참조)에서도 정신건강과 관련된 다양한 결과 지표를 제공한다. 그러나 이들 중 일부는 비정기적으로 제공되거나 일부 OECD 국가에 한하여 제공되며, 경우에 따라 비공식적인 출처를 기반으로 작성되었다는 한계가 있다².

상자 1.1. 사람들의 웰빙 측정

OECD 웰빙 프레임워크는 인적, 사회적 조건을 측정하고 OECD 국가에서 거주하는 사람들의 삶이 전반적으로 개선되고 있는지를 평가하기 위한 광범위한 결과 중심 측정 도구이다(OECD, 2020_[4]). 여기에는 개인, 가정 및 공동체 수준의 생활 조건을 ‘지금 여기(here and now)’에서의 삶에 초점을 맞춘 현재의 웰빙과 미래의 웰빙을 지속시키기 위해 필요한 체계적인 자원이 포함된다³. 웰빙 프레임워크는 OECD *How's Life?* 보고서 시리즈와 웰빙 관련 기타 다양한 OECD의 노력을 뒷받침하는 근거가 된다(개요는 <https://www.oecd.org/wise/> 참조).

그림 1.1. OECD 웰빙 프레임워크



출처: OECD(2020_[4]), *How's Life? 2020: Measuring Well-being*, OECD Publishing, Paris, <https://doi.org/10.1787/23089679>

정신건강은 웰빙 프레임워크 내에서 별도 영역으로 명시적으로 규정되어 있지 않지만, 정신건강 결과는 다음과 같이 여러 영역과 관련이 있다.

- 첫째로 프레임워크에서 광범위한 ‘건강’ 영역은 정신건강과 신체건강을 모두 포괄한다는 점이 가장 중요하다. 예를 들어 정신건강 문제에 관한 두 가지 지표(자살, 급성 알코올 남용 및 약물 과다복용으로 인한 사망, 우울증 발병 위험이 있는 사람의 비율)가 OECD *How's Life? 2020* 보고서의 건강 영역에 포함되었다.
- 둘째, ‘주관적 웰빙’ 영역은 긍정적인 심리적 기능의 요소, 특히 행복과 긍정적 및 부정적 정서를 포함한다. 자신의 삶에 대한 평가(예: 삶의 만족도)도 이 영역에 포함된다.
- 마지막으로 미래의 웰빙을 위한 자원에 포함된 ‘인적 자본’은 개인의 지식, 역량, 기술 및 건강 상태를 의미하며, 미래의 웰빙에 대한 기여의 측면에서 이 영역에 해당하고 여기에는 조기 사망률, 비만 유병률 등의 지표가 포함된다. (OECD, 2020_[4])⁴.

또한 사회적 관계, 재정적 안정(소득과 부), 지식과 역량 등 광범위한 정의의 긍정적 정신건강에 포함되는 여러 긍정적인 기능은 OECD 웰빙 프레임워크 내에서 별도의 영역에 포함된다.

본 보고서의 목적은 공식 데이터 생산기관에서 인구 정신건강에 대한 데이터 조사를 수행하는 빈도를 높이고 국제적으로 통일성 있는 방식으로 데이터를 수집하도록 장려하는 것이다. 이는 궁극적으로 개인이 자신과 자신이 선택한 삶을 살 수 있는 능력(이 경우 정신적으로 건강하다고 느끼고 정신적 고통이 없는)에 있어 중요한 요인을 평가하는 웰빙 접근법의 원칙과 일치한다(OECD, 2020^[4]). 또한 빈번하게 사용하는 투입 지표 또는 산출 지표로 측정된 웰빙의 여러 동인이 해당 결과와 불완전한 상관관계가 있을 수 있다(예: 의료 시스템을 이용할 수 없는 경우 정신건강 지출이 정신건강 상태를 제대로 반영하지 못하며, 이와 마찬가지로 약물 처방 횟수로 사람들의 정신건강 상태를 파악하기 어려움)(OECD, 2011^[8]).

의료 전문가에게 진단을 받거나 치료를 받은 대상만이 아니라 전체 인구를 대상으로 정신건강 상태에 대한 데이터를 수집하는 것은 여러모로 중요하다. 첫째, 진단받은 사람 수에 초점을 맞춘 측정은 국가의 의료 시스템에 대한 접근성과 발전 정도, 그리고 사람들(및 특정 인구 집단)이 치료를 받을 가능성 등에 영향을 받는다는 한계가 있다. 둘째, 정신건강 문제의 예방 전략은 정신건강 위험군을 조기에 파악하는 데 도움이 된다. 따라서 의료 시스템 이용 전과 후의 결과를 추적해야 한다. 셋째, 긍정적 정신건강은 인구의 기초 자산이므로 그 자체로 추적할 가치가 있다. 일반적으로 인구(조사) 통계에서 수집되는 사람들의 물질적 조건, 삶의 질, 사회적 관계(및 이러한 요인의 불평등)와 같은 광범위한 위험 및 회복탄력성 요인을 정신건강과 연계한다면 정신건강 전략을 수립하는 데 도움이 될 수 있다.

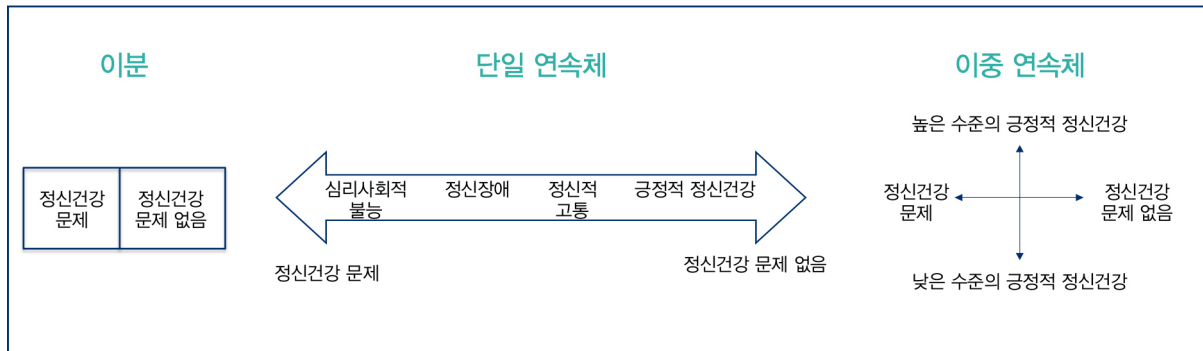
정신건강의 개념: 질병에서 건강까지

이전 OECD 연구에서 정신건강의 정의는 세계보건기구(WHO)에 의해 널리 인정되는 정의를 채택했으며 구체적인 정의는 다음과 같다. ‘개인이 자신의 능력을 실현하고 일상적인 삶의 스트레스에 대처할 수 있으며 생산적으로 일할 수 있고 자신의 공동체에 긍정적인 기여를 할 수 있는 웰빙 상태’(OECD, 2021^[3], OECD, 2015^[6]). 이 정의는 정신건강이 질병이 없는 상태만을 의미하는 것이 아니며, 심리사회적 기능의 다양한 측면을 포괄하는 개념임을 명시하고 있다(세계보건기구, 1948^[9])⁵.

지난 수십 년 동안 정신건강에 관한 다양한 이론이 개발되었다. 정신질환의 증상 유무에 초점을 맞춘 이론(‘이분 모델’)부터 정신건강을 경험의 스펙트럼으로 간주하는 이론(‘단일 연속체 모델’), 정신건강 문제와 긍정적 정신건강이 서로 연관되어 있지만, 별개의 경험으로 보는 이론(‘이중 연속체 모델’)에 이르기까지 다양하다(그림 1.2). 각 모델은 서로 다른 함의를 내포하고 있어 정신건강의 개념을 전체적으로 파악하기 위해서는 정신건강 결과를 상기 순서대로 추적해야 한다.

그림 1.2. 정신건강의 모델

정신건강의 개념적 프레임워크



출처: lasiello 외 (2020_[10]), “Mental Health and/or Mental Illness: A Scoping Review of the Evidence and Implications of the Dual-Continua Model of Mental Health”, Evidence Base, 10.21307/eb-2020-001.; Keyes, C. (2005_[11]), “Mental illness and/or mental health? Investigating axioms of the complete state model of health”, Journal of Consulting and Clinical Psychology, 73(3): 539에서 인용

이분 모델

임상심리학 및 정신의학과 연구는 역사적으로 정신건강을 개선하기 위해 정신질환 증상, 즉 정신병리를 줄이는 데 초점을 맞춰왔다. 이러한 ‘질병 중심’ 관점에서는 정신질환(정신의학과 분류 체계에 의해 정의된 상태)이 중심 개념이며, 치료의 목표는 건강하고 행복한 삶으로 회복하도록 지원하는 것이 아니라 주로 관련 증상을 줄이는 데 있다. 이러한 관점에서 개인은 두 가지 상태 중 하나를 경험할 수 있는데, 이는 정신질환으로 진단받거나 정신적으로 건강하다고 가정하는 것이다(Routledge 외, 2016_[12], Keyes, 2005_[11], Trent, 1992_[13]).

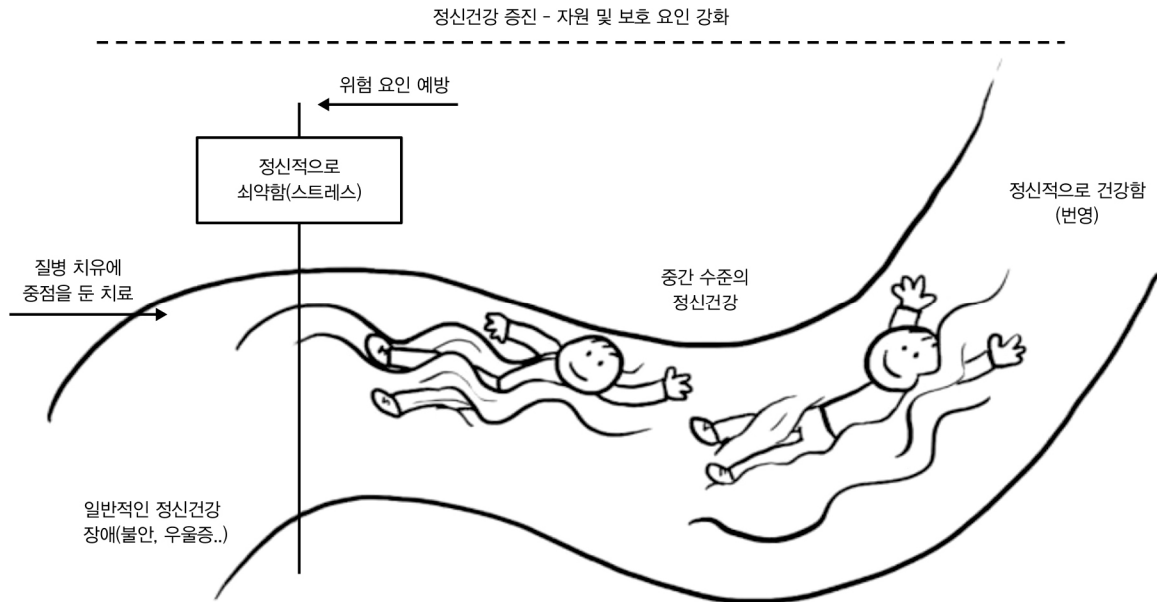
예를 들어 정신질환의 이분법적 분류는 적절한 의료 서비스나 기타 지원 서비스를 이용하거나 진단된 상태를 관리하기 위한 지침 및 치료 경로를 정의하는 데 유용할 수 있다. 그러나 의료진과 연구자는 이분 모델의 환원주의적 특성, 즉 질병이 완전히 건강한 상태로 전환되는 임의의 지점에 대한 개념과 진단이 불가능한 상태의 역치를 넘으면 정신적으로 건강한 것이라는 가정에 대해 비판해 왔다(Herron 및 Trent, 2000_[14]).

연속체로서의 정신건강

대안적 접근법은 정신건강을 스펙트럼의 한쪽 끝에 위치한 심각한 정신건강 문제부터 다른 쪽 끝에 위치한 긍정적 정신건강(높은 수준의 정서적 및 심리적 웰빙)에 이르는 연속적인 경험으로 특징짓는다(Patel 외, 2018_[15], Payton, 2009_[16], Greenblat, 2000_[17]). 이러한 관점은 위험, 증상 및 문제에 초점을 맞추는 전통적인 방식을 넘어 건강과 웰빙을 개선하는 요인에 초점을 맞추는 ‘건강생성(salutogenic)’ 접근법, 즉 사람의 행복과 건강을 유발할 수 있는 접근법에 기초를 둔다. 이 접근법은 사람들의 다양한 경험을 인정한다(단순히 걱정이 많거나 수면에 어려움을 겪는 사람과 본격적으로 주요우울삽화를 경험하는 사람의 경험은 다를 수 있음).

이 모델에서는 ‘모든 사람은 정신건강을 가지고 있으며’ 개인은 자신이 처한 상황, 직면한 문제 및 이용할 수 있는 내외부 자원에 따라 일생에 걸쳐(매일 또는 매주까지 포함) 스펙트럼의 위아래로 이동할 수 있다. 일부 연구자들은 극심한 정신건강 문제와 긍정적 정신건강 사이에서 다양한 상태를 가변적으로 끊임없이 겪는 과정을 표현하기 위해 선형적인 연속체가 아닌 강이라는 은유적인 표현을 사용했다(그림 1.3)(Koushede 및 Donovan, 2022_[18]).

그림 1.3. 강과 같은 정신건강 연속체



출처: Koushede, V. 및 R. Donovan (2022_[18]), “Applying Salutogenesis in Community-Wide Mental Health Promotion”, *The Handbook of Salutogenesis*. Springer, Cham. https://doi.org/10.1007/978-3-030-79515-3_44.

정책적 관점에서 정신건강 문제와 긍정적 정신건강 사이의 전체 스펙트럼을 고려하는 것은 개입 목표를 설정하고 설계하는 데 모두 영향을 미치며, 사후대응이 아닌 사전예방적 시스템 및 서비스 설계를 위한 새로운 길을 제시할 수 있다. 단일 연속체 모델은 이분법적 관점에 비해 정신건강 전략이 (진단된) 질병을 ‘치유’하거나 관련 증상을 줄이는 데 그치는 것이 아니라, 스펙트럼의 중간에 위치한 사람들의 상태가 악화되는 것을 방지하고 정신적으로 건강한 인구를 확대하는 데 초점을 맞출 수 있도록 하는 가치를 추가했다.

이중 연속체로서의 부정적 및 긍정적 정신건강

세 번째 개념적 관점은 세계보건기구(WHO), 공중보건기관, 국가 통계청 및 기타 정부부처 등 여러 국제적 참여자들의 관심이 확대되고 있으며, 정신건강 문제와 긍정적 정신건강을 더욱 명확하게 구분한다 (Statistics Canada, n.d._[19], Australian Early Development Census, 2012_[20], Swiss Health Observatory, n.d._[21], Government of Western Australia Mental Health Commission, 2021_[22], Queensland Government, 2015_[23]) (세계보건기구, 2022_[24]). ‘이중 연속체’ 모델은 정신건강 문제와 긍정적 정신건강 상태가 단일 스펙트럼의 양극단에 위치하는 것이 아니라, 서로 연관되어 있지만 별개의 경험(두 개의 서로 다르지만 교차하는 연속체에 위치)으로 특징짓는다⁶.

정신건강 문제와 긍정적 정신건강 또는 높은 수준의 정서적, 심리적 웰빙은 서로 밀접하게 연결되어 있다. 인구 수준에서 정신건강이 개선되면 시간이 경과함에 따라 정신장애 평균이 감소함을 의미하는 반면, 긍정적인 정신건강을 경험하면 정신장애가 발생할 위험이 감소하고 정신장애가 발생한 경우에도 회복을 돕는 중요한 회복력 요인으로 간주된다(Keyes, Dhingra 및 Simoes, 2010_[25], Robinson, 2012_[26], Santini 외, 2022_[27]).

그러나 이중 연속체 모델을 지지하는 사람들은 단일 연속체 모델에서 제안하는 바와 같이 정신건강 문제와 긍정적 정신건강 사이의 연관성은 선형적이지 않으며, 임상적으로 유의미한 정신질환의 증상이 발현되거나 질환을 진단받은 적이 없다는 점이 항상 건강한 정신 상태를 의미하지는 않는다고 주장한다.

반대로 정신장애 증상과 이와 관련된 고통 및 장애를 가지고 있지만, 자신의 삶에 만족하고 잠재력을 발휘하는 사람도 있을 수 있다(Galderisi 외, 2015_[28]). 이중 연속체 모델은 인간 경험의 다양성을 완전히 인정하는 것을 목표로 하며, 이러한 관점은 정신질환을 실제 경험한 사람들이 공감하는 경우가 많다(New Zealand Initial Mental Health and Wellbeing Commission, 2020_[29]).

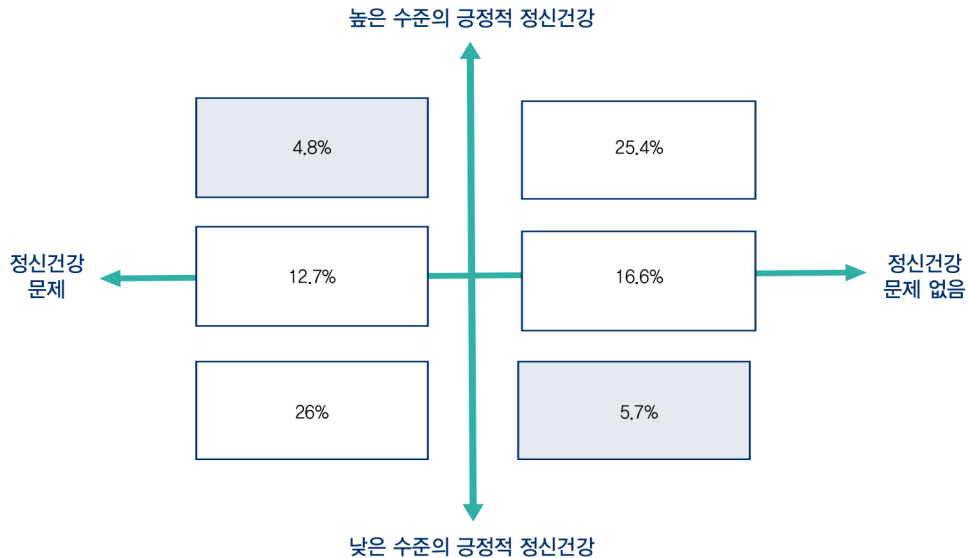
이중 연속체에 대한 대부분의 연구에서는 설문조사 데이터가 단일 또는 이중 연속체 모델 중 어떤 모델에 가장 적합한지 비교하기 위해 확인적 요인분석(CFA)에 의존해 왔다. Keyes(2005_[11])는 미국 전역을 대표하는 성인 표본을 대상으로 정서적, 심리적, 사회적 웰빙 측면⁷과 몇 가지 일반적인 형태의 정신질환(지난 1년간 주요우울삽화, 범불안장애, 공황장애 또는 알코올 의존증 유무)을 측정했다. 그 다음 두 요인이 서로 상관이 있지만 별개의 잠재 변인이 존재한다는 것을 강조하기 위해 확인적 요인분석을 실시했다. 미국 이외의 인구를 대상으로 긍정적 정신건강과 정신질환에 대해 다양한 측정 도구를 사용하여 진행된 추가 연구도 이중 연속체 모델의 개념을 추가로 뒷받침했다. 최근 검토에 따르면 횡단 연구, 종단 연구, 중재 연구를 포함하여 동료 검토가 수행된 83편의 실증적 논문이 확인되었으며, 이 논문들은 단일 연속체 모델에 비해 이중 연속체 모델의 우수한 설명력을 뒷받침했다(Iasiello, van Agteren 및 Cochrane, 2020_[10], Routledge 외, 2016_[12])⁸.

그러나 이중 연속체 모델에서 일반적으로 두 축이 완전히 직교하는 시각화는 오해의 소지가 있을 수 있다. 여러 연구에서 ‘완전함’(정신질환 없음, 높은 수준의 긍정적 정신건강), ‘취약함’(낮은 수준의 정신질환, 낮은 수준의 긍정적 정신건강), ‘증상은 있으나 만족함’(높은 수준의 정신질환, 높은 수준의 긍정적 정신건강), ‘심각함’(높은 수준의 정신질환, 낮은 수준의 긍정적 정신건강) 등 다양한 범주를 사용하여 개인을 모델의 사분면에 따라 별도의 그룹으로 분류한다(Iasiello, van Agteren 및 Cochrane, 2020_[10]). 그러나 이러한 범주 내의 분포는 긍정적 정신건강과 정신건강 문제의 수준 간에 매우 밀접한 관련이 있으며, 정신건강 상태가 정서적, 심리적 웰빙에 심각한 손상을 초래한다는 점을 강력하게 시사한다. 예를 들어 호주 초등학생을 대상으로 한 연구에 따르면 약 5%의 아동만이 높은 수준의 긍정적 정신건강을 경험하면서 정신건강 문제도 경험하거나, 낮은 수준의 긍정적 정신건강을 경험하면서 정신건강 문제를 경험하지 않는 것으로 나타났다(그림 1.4). 이와 유사하게 Keyes의 연구 결과에 따르면 지난 1년간 정신질환 진단을 받지 않은 응답자 5명 중 1명만이 높은 수준의 긍정적 정신건강을 경험한 것으로 보고되었으나 정신장애를 가진 응답자의 경우 그보다 훨씬 적은 수준으로 나타났다(표 1.1)(Keyes, 2005_[11]).

또한 동일한 연구에서 긍정적 정신건강에 대한 경험은 지난 1년간 경험한 심리적 장애의 유형(및 조사 시점의 심각도)에 따라 크게 달라졌는데, 범불안장애가 있는 응답자의 경우 2%에 불과한 반면 알코올 의존증이 있는 응답자는 8%에 달하는 등 다양하게 나타났다(표 1.1). 그럼에도 불구하고, 정신건강 상태가 심각한 수준인 응답자 중 일정 수준의 긍정적 정신건강 상태에 도달할 수 있는 응답자의 비율이 적지 않은 것으로 나타났다. 잘 알려지지 않았지만 정서적, 심리적 웰빙의 다양한 측면과 심각한 정신건강 상태 사이에서도 상호 보완이 가능하다. 일부 연구에 따르면 심리적으로 건강한 성인에 비해 우울증 환자는 부정적인 사건에 둔감하게 반응한다. 양극성장애 환자는 조증이 발생하는 동안 긍정적인 감정을 더 크게 느끼고 조현병 환자는 환각이나 망상으로부터 의미를 구성할 수 있으며, 외상 생존자는 고통스러운 경험을 다루면서 의미있는 삶을 살 수 있다(Goodman, Doorley 및 Kashdan, 2018_[30]).

그림 1.4. 호주 아동의 정신건강의 이중 연속체

정규 초등교육과정 1학년 학생 중 다양한 정도의 정신건강 문제(불안, 우울, 행동 문제)와 긍정적 정신건강(심리사회적 기능)을 경험하는 아동의 비율(호주, 2012)



주: 데이터는 2012년에 전국적으로 실시된 호주인의 조기발달지수(Australian Early Development Index)에서 추출되었으며 아동의 정신건강에 대한 응답은 해당 학교 교사가 제공하였다. 어두운 음영으로 표시된 부분은 긍정적 정신건강은 낮은 수준이지만 정신질환을 경험하지 않는 아동 또는 정신질환을 경험하지만 긍정적 정신건강이 높은 수준인 아동의 비율을 나타낸다. 원출처에서는 해당 범주를 정신건강 어려움(예: 불안장애, 우울증, 행동 문제) 및 정신건강 역량(예: 건강한 심리사회적 기능)으로 지칭한다.

출처: Australian Early Development Census (2012_[20]), *The mental health of Australian children: A dual continuum*, <https://www.aecd.gov.au/resources/detail/the-mental-health-of-australian-children-a-dual-continuum>.

표 1.1. 긍정적 정신건강 달성 정도에 영향을 미치는 정신건강 상태의 차이

지난 12개월 동안 정신건강 어려움을 겪고 있는 응답자 중 낮은 수준, 중간 수준, 높은 수준에서 긍정적 정신건강의 경험을 보고한 성인 응답자의 비율(미국, 1994~1995)

| | 낮은 수준의 긍정적 정신건강 (쇠약함) | 중간 수준의 긍정적 정신건강 | 높은 수준의 긍정적 정신건강 (번영) |
|------------------|-----------------------|-----------------|----------------------|
| 전체 표본 (n=3,032명) | 16.9% | 65.1% | 18% |
| 주요우울삽화 (n=422명) | 33.9% | 60.2% | 6.2% |
| 범불안장애 (n=98명) | 55.1% | 41.8% | 2% |
| 공황장애 (n=204명) | 39.2% | 58.3% | 2.5% |
| 알코올 의존증 (n=194명) | 24.7% | 69.1% | 7.7% |
| 중복이환 (n=193명) | 43.5% | 55.4% | 2.1% |

주: 데이터는 ‘Midlife in the United States’ 연구에서 추출되었다. 정신장애는 단축형 복합 국제진단면담도구(Composite International Diagnostic Interview Short Form, CIDI-SF)로 측정하였다. 번영(쇠약함)은 긍정적 정서에 관한 두 가지 문항 중 하나에서 높은(낮은) 수준을, 긍정적 기능에 관한 11개의 문항 중 6개 문항에서 높은(낮은) 수준을 보이는 개인으로 정의하였다(Ryff의 심리적 웰빙 척도 및 Keyes의 사회적 웰빙 척도에 따라). 그 외의 모든 응답자는 중간 수준의 긍정적 정신건강으로 분류되었다. 동반이환 조합에 관계없이 두 가지 이상의 정신건강 상태를 경험하는 것을 의미한다.

출처: Keyes, C. (2005^[11]), “Mental illness and/or mental health? Investigating axioms of the complete state model of health”, *Journal of Consulting and Clinical Psychology*, 73(3): 539.

이중 연속체 모델은 정신건강 문제와 긍정적 정신건강을 구분함으로써 각 동인의 상대적 중요성이 다르다는 것을 암묵적으로 시사한다. 정신질환을 예방하고 긍정적 정신건강을 향상하는 데 동일한 전략을 적용하는 것은 적절하지 않을 수 있으므로 이중 연속체 모델은 정책 수립 및 임상 실무에 중요하다. 이를 뒷받침하는 근거가 계속 발견되고 있다. 덴마크와 영국의 일부 인구기반 연구에 따르면 사람들의 물질적 조건과 삶의 질(예: 저소득 및 교육 수준, 일자리 부족 및 재정적 불안정)의 박탈은 각 연속체의 끝단에서 결과를 예측하는 것으로 나타났다(즉 정신건강 문제 및 낮은 수준의 긍정적 정신건강 모두). 이러한 사회경제적 요인은 높은 수준의 긍정적 정신건강을 결정하는 데에는 중요한 역할을 하지 않는 것으로 나타났다(Stewart-Brown 외, 2015^[31], Nielsen 외, 2016^[32], Santini 외, 2020^[33]). 그러나 캐나다와 슬로베니아의 인구기반 데이터에 따르면 재정적 안정과 가계 소득이 높을수록 실제로 심리적 번영감이 높아지는 것으로 나타났다(Varin 외, 2020^[34], Vinko 외, 2022^[35]). 반면 대부분의 연구에 따르면 사회적 유대감 강화, 가족 관계 개선 및 여가 활동 참여 등 관계적 요인은 정신건강 상태의 위험을 감소시키고 긍정적 정신건강을 향상하는 데 관련성이 있는 것으로 나타났다(Van Lente 외, 2012^[36], Santini 외, 2020^[33], Santini 외, 2017^[37], Solin 외, 2019^[38], Thoits, 2011^[39]).

이중 연속체 모델에 추가된 가치는(단일 연속체에 비해) 정신건강 문제와 긍정적 정신건강 상태가 사람들의 삶에 영향을 미친다는 점을 더욱 명확하게 전달한다는 것이다. 측정 관점에서 이중 연속체 모델은 인구 조사 및 건강 평가에서 정신건강과 긍정적 기능에 대한 데이터를 수집하면 정신건강에 대한 이해도를 높일 수 있음을 시사한다. 아울러 정신건강의 개선 및 정신질환 감소라는 두 가지 목표와 관련된 요인을 파악하고 더 나아가 정책 수단을 파악하는 데 도움이 될 것이다. 따라서 본 보고서에서는 정신건강 문제와 긍정적 정신건강을 개별적으로 고려하고 다음 섹션에서 각각에 대해 자세히 정의한다.

정신건강 문제

정신건강 문제라는 용어는 진단 가능한 정신적, 행동적 상태는 물론 심리적 고통이라는 초진단적 특성을 의미하기도 한다.

정신건강 상태

본 보고서에서 사용된 ‘상태’ 또는 ‘장애’라는 용어는 세계보건기구 국제질병분류(ICD) 또는 미국 정신 의학협회 정신질환 진단 및 통계 편람(DSM)과 같은 정신의학과 분류 체계에 따라 임상적 진단 기준에 도달하는 증상을 설명하기 위해 사용된다⁹. 이러한 분류 체계에는 경도 또는 중등도 불안 및 우울증, 약물 및 알코올 사용 장애, 중증 우울증, 양극성장애, 조현병과 같은 중증 장애를 포함하여 100개 이상의 개별 진단 및 장애가 수록되어 있으며 각각 고유한 증상, 발병 연령 및 궤적이 수록되어 있다(상자 1.2). 정신건강 상태의 경험은 일생에 걸쳐, 또는 몇 주의 짧은 기간에 걸쳐 매우 가변적일 수 있다. 예를 들어 중등도의 우울삽화를 경험하는 개인의 정신건강 상태가 심각한 수준으로 악화될 수 있으며, 이와 마찬가지로 중증의 우울삽화 증상이 완화되어 정신건강 상태가 안정될 수 있다(OECD, 2021^[3]).

인구의 절반이 일생동안 한 번의 정신건강 상태를 경험하고 매년 5명 중 1명이 정신건강 상태를 경험하는 것으로 추정된다(OECD, 2019_[40]). 현재 인구기반 조사 데이터는 불안과 우울증 증상 경험에 초점을 맞추고 있는 경우가 많다(제2장 참조). 보건계량연구소(Institute for Health Metrics and Evaluation, IHME)에서 코로나19 이전에 조사한 추정치에 따르면 유럽연합(EU) 국가에서 가장 일반적으로 발생하는 정신장애는 불안장애로 2016년에 약 2,500만 명(전체 인구의 5.4%)이 이 질환을 앓고 있는 것으로 나타났으며, 약 2,100만명(전체 인구의 4.5%) 이상이 우울장애를 앓고 있는 것으로 나타났다. EU 국가 전체에서 약 1,100만 명(2.4%)이 약물 및 알코올 사용 장애를 앓고 있는 것으로 추정된다. 양극성장애와 같은 중증 정신질환은 약 500만 명(전체 인구의 1%)이, 조현병은 150만 명(0.3%)이 앓고 있는 것으로 나타났다(OECD/European Union, 2018_[41])¹⁰.

상자 1.2. 정신건강 상태 및 증상의 예

정신질환 진단 및 통계 편람(DSM)에 따르면 정신건강 상태는 개인의 인지, 감정 조절 또는 행동에 임상적으로 심각한 어려움과 장애를 특징으로 하는 증후군으로, 이는 정신적 기능의 근간이 되는 심리적, 생물학적 또는 발달적 과정에 나타나는 기능 장애를 반영하고 있다. 정신장애는 일반적으로 사회적, 직업적 또는 기타 중요한 활동에서 심각한 고통이나 장애와 관련이 있다(American Psychiatric Association, 2013_[42]). 정신장애는 신체적 질환을 동반하거나 정신건강 문제가 복합적으로 나타나는 것이 일반적이다. 가장 최신 버전인 정신질환 진단 및 통계 편람의 5번째 개정판(DSM-5)은 2013년에 발간되었으며, 총 157개의 진단과 300여 개의 장애가 수록되어 있다. 가장 일반적인 장애군은 다음과 같다.

기분/정동장애

기분장애 또는 정동장애는 개인의 기능에 지장을 주는 일반적인 감정적 상태의 장애를 특징으로 한다. 다양한 형태의 기분장애가 존재하며, 예를 들어 주요우울장애는 기분저하, 자존감 저하 및 평소 즐기던 활동에 대한 흥미 상실의 최소 2주 이상 지속되는 것이 특징이다. 피로, 두통 또는 소화장애 등의 신체적 증상도 흔히 나타난다. 양극성장애는 우울증과 조증(병으로 고양된 기분, 각성 및 에너지 수준)이 번갈아 나타나는 것이 특징이다.

불안장애

불안장애는 불안과 두려움을 과도하게 느껴 통제할 수 없는 느낌이 드는 것이 특징이다. 구체적인 증상은 불안장애의 유형에 따라 다르다. 가장 일반적인 불안장애는 범불안장애, 공황장애 및 사회불안장애이다. 또한 고소공포증이나 비행공포증과 같은 특정 공포증(특정한 대상이나 상황에 대한 두려움)도 다양하게 존재한다.

약물사용장애

약물사용장애는 부정적인 결과에도 불구하고 통제할 수 없이 약물을 복용하는 것이 특징인 질환으로 정서적, 신체적, 행동적 문제를 동반하며, 여러 번 시도했음에도 불구하고 약물 복용을 중단하지 못하는 경우가 많다. 예를 들어 알코올사용장애는 약물남용장애의 한 유형으로 잦은 음주와 폭음을 포함한다.

적응장애

적응장애는 심리사회적 스트레스 요인에 대한 부적응적인 정서적 또는 행동적 반응이 특징이다. 적응장애는 개인이 일상생활에서 겪는 여러 가지 스트레스가 높은 생활 사건에 적응하거나 대처하는 데

상당한 어려움을 겪을 때 발생한다.

예를 들어 외상 후 스트레스 장애(PTSD)는 일반적으로 전쟁, 성폭행 또는 아동 학대와 같은 충격적인 생활 사건이나 위협적인 상황에 노출되어 발생한다. 수면장애, 집중력 저하 또는 과민성부터 과각성 및 과장된 놀람 반응에 이르기까지 다양한 증상이 나타날 수 있다.

정신병적 장애

정신병적 장애는 망상과 환각을 일반적인 증상으로 동반하는 심각한 정신건강 상태를 말한다. 가장 일반적인 정신병적 장애는 조현병으로 현실을 비정상적으로 해석하고 경험하며 환각, 망상, 와해된 언어와 행동 등이 복합적으로 나타나 일상적인 기능을 손상시키는 것이 특징이다.

성격장애

인격장애는 문화적·사회적 규범과 크게 다른 행동, 인지 및 내적 경험의 부적응 패턴이 만연하며 오래 지속되는 상태를 말한다. 이는 인지, 정서, 대인관계 기능 또는 충동 조절의 어려움과 관련이 있다. 인격장애에는 이상하고 기이한 장애, 극적이고 감정적인 장애, 불안하고 두려워하는 장애 등 세가지 범주의 성격장애가 있다.

신체형 장애 및 해리장애

신체형 장애는 심리적 장애가 신체적 증상으로 나타나는 질환을 말한다. 해리장애는 기억, 인식, 지각 또는 정체감 문제가 포함되며, 해리 장애가 있는 사람은 자신의 몸이 분리된 느낌을 느끼거나 다른 정체감을 발달시킬 수 있다.

섭식장애

섭식장애는 신체 또는/그리고 정신건강에 영향을 미치는 비정상적인 섭식행동이 특징이다. 다양한 유형의 섭식장애가 존재하며 신경성 폭식증, 신경성 식욕부진 및 폭식장애가 가장 일반적이다. 섭식장애는 불안 장애, 우울증 및 약물남용과 동반되어 나타나는 경우가 많다.

강박장애

강박장애는 불안감, 혐오감 또는 불편감을 유발하는 침투적이고 반복적인 생각이나 정신적 이미지(강박사고)와 이러한 불편감을 해소하기 위해 손씻기, 숫자 세기, 청소 또는 정리 정돈과 같은 특정한 작업이나 일과(강박행동)를 수행하고자 하는 충동이 특징인 정신 및 행동 장애이다.

출처: (세계보건기구, 2021^[43], American Psychiatric Association, 2013^[42]).

심리적 고통

본 보고서에서 ‘심리적 고통’이라는 용어는 정신의학적인 분류 체계 내에서 임상적 진단 기준에 도달하지 않지만, 종종 신체적 증상(예: 불면 또는 식욕 저하)이 동반하는 부정적 정서의 비특이적 증상(예: 슬픔, 괴로움, 안절부절못함)을 지칭하기 위해 사용된다.

심리적 고통과 정신건강 상태가 개념적으로 구별되는지 여부에 대해 논쟁이 있다. 일부 연구자들은 심리적 고통이 지속적이고 부정적 사건이나 다른 스트레스 요인에 대한 ‘예상 가능한 반응’을 초월하는 경우에만 병리적 심리적 과정의 일부이자 정신건강 상태의 지표로 간주해야 한다는 점에서 질적으로 구별된다고 주장한다. 그러나 이는 실제로 판단하기 어렵고 개인의 전반적인 삶 또는 사회경제적 조건에 따라

달라질 수 있다(Horwitz, 2007^[44], Phillips, 2009^[45], Payton, 2009^[16], Roger T. Mulder, 2008^[46], Wakefield 외, 2007^[47]). 정신질환 진단 및 통계 편람 제5판(DSM-5)은 정신적 고통이 임상적으로 유의한지를 판단하는 기준을 제공하지 않으며, 일반적으로 ‘적절성’보다는 정신적 고통으로 인한 기능의 손상 정도에 따라 평가한다.

본 보고서의 제2장에 설명된 바와 같이 개인의 심리적 고통을 평가하기 위해 개발된 도구를 사용하면 정신건강 상태의 심각한 사례와 심각하지 않은 사례를 명확히 구분할 수 있다. 이는 대부분의 정신건강 상태의 초진단적 특성인 정신장애와 정신적 고통이 실제로 밀접하게 관련되어 있음을 시사한다(Barlow 및 Durand, 2009^[48]). 더욱이 심리적 고통의 경험이 일시적일지라도 개인에게 심각한 고통과 어려움을 의미할 수 있으므로 그 자체로 주목할 가치가 있다.

긍정적 정신건강

긍정적 정신건강은 심리적, 정서적, 경우에 따라 사회적, 관계적, 영적 웰빙을 포괄한다(Huppert, 2005^[49], Keyes, 2005^[11], Steger 외, 2006^[50], Reis 및 Gable, 2003^[51])¹¹.

긍정적 정신건강의 개념은 주관적 웰빙의 개념과 밀접한 관련이 있으며, 이는 ‘사람들이 자신의 삶에 대해 긍정적 또는 부정적인 다양한 평가를 내리고 자신의 경험에 대해 정서적 반응을 보이는 양호한 정신 상태’를 의미한다(OECD, 2013^[52]). 2013년 OECD는 *Guidelines on Measuring Subjective Well-being(주관적 웰빙 측정 지침)*을 발표했으며, 이는 다음과 같은 주관적 웰빙의 세 가지 광범위한 측면을 식별한다(그리고 데이터 수집기관을 위한 척도를 제안한다).

- **생애 평가(Life evaluation)** - 개인의 삶 전체 또는 특정 측면에 대한 반성적 자기평가(예: 삶의 만족도 측정, 재정 상황에 대한 만족도).
- **정서(Affect)** - 일반적으로 특정 시점을 기준으로 측정되는 개인의 느낌, 감정 또는 상태(예: 행복, 걱정, 고통, 피로의 경험에 대한 측정).
- **행복(Eudaimonia)** - 삶의 의미와 목적 또는 좋은 심리적 기능(예: 삶에서 하는 일이 가치 있다고 느끼는지에 대한 측정).

긍정적 정신건강과 주관적 웰빙에서 가장 많이 중첩되는 영역은 정서(일반적인 정신건강 측정에서 걱정, 고통 또는 피로와 같은 특정 정서 상태의 지속적인 경험을 강조하는 경우)와 행복(긍정적 정신건강을 포착하기 위해 개발된 많은 척도에서 명백하게 드러나는)이다. 긍정적 정신건강 측정에 부차적인 개념인 자율성, 낙관주의, 회복력 또는 환경 통제력 등은 전술된 주관적 웰빙에 관한 OECD의 정의에서 명시적으로 언급되지 않는다(Davydov 외, 2010^[53], Snow, 2019^[54], Peterson 및 Seligman, 2004^[55], Conversano 외, 2010^[56], Ryff 및 Keyes, 1995^[57]). 해당 개념은 앞서 언급한 *OECD 지침*(부속서 1 및 모듈 D 참조)에서 논의된 행복 및 심리적 기능에 대한 일부 척도(확장형)에 포함되지만, 낙관주의 및 기타 성격 특성과 같은 평가 방식은 주관적 웰빙에 대한 최종 결과보다 생활 환경에 대한 개인의 정서적 반응에 영향을 미치는 매개요인으로 간주된다(OECD, 2013^[52]).

주관적 웰빙의 개념과 긍정적 정신건강의 개념 간의 가장 큰 차이는 한 개인의 정신건강만을 평가하는 것이 아니라, 한 개인의 삶에 대해 모든 영역에서 광범위하게 평가가 이루어지는 생애 평가 측정에 관한 것이다. 그럼에도 불구하고, 생애 평가 측정은 개인의 웰빙에 대한 인식을 전반적으로 반영하는 광범위한 결과 측정으로서 가치가 있으므로 실제로 (긍정적) 정신건강에 대한 연구에 포함되는 경우가 많다.

제2장에서는 앞서 언급한 주관적 웰빙의 세 가지 측면과 정서적, 심리적, 사회적 웰빙 측면을 포괄하는 긍정적 정신건강 척도(주로 긍정심리학에서 파생됨)에 대해 현재 OECD 국가의 데이터 수집 관행을 검토한다.

제3장에서는 정신건강 도구의 통계 품질을 논의한다. *OECD 지침*에서 이미 생애 평가, 정서 및 행복을 측정하는 문제를 심도 있게 다루었기 때문에 후자에만 중점을 두고 논의한다(OECD, 2013_[52]). 특히 정서와 행복을 측정하는 주제는 향후 주관적 웰빙에 대한 OECD 업무에서도 지속적으로 탐구될 것이다. 본 간행물에서 신체적 및 성적 건강, 재정적 안정 또는 학업 및 직업적 성과(OECD 웰빙 프레임워크의 다른 부문에서 다루고 있음)와 같은 영역을 지나치게 광범위하게 포함하는 긍정적 정신건강에 대한 정의는 고려하지 않았다 (Fusar-Poli 및 Santini, 2022_[58], Fusar-Poli 외, 2020_[59], Harvard Center for Health and Happiness, n.d._[60]).

결론

사람들의 삶에 중요한 웰빙 결과를 충분히 평가하려면 정신건강을 측정하는 것이 중요하다. 본 보고서의 목적은 공식 데이터 생산기관이 정신건강 상태에 대한 인구 수준의 데이터를 조사하는 빈도를 높이고 보다 국제적으로 통일성 있는 방식으로 데이터를 수집하도록 장려하여 의료 시스템을 이용하는 대상뿐만 아니라 모든 사회 집단의 정신건강에 대해 이해하고, 오늘날의 사회와 경제에서 공공 정책 과제로 점차 인식되고 있는 문제를 해결하기 위한 것이다.

정신건강은 정신질환의 유무를 이분법적으로 구분하는 것 이상의 다면적인 개념이다. 정신건강의 모든 측면을 고려하면 정신건강 시스템과 서비스를 사후대응이 아닌 사전예방적으로 설계할 수 있는 새로운 방법을 제시하고, 긍정적 정신건강에 대한 관심 그 자체의 중요성에 주목하며, 질병을 줄이고 긍정적 정신건강을 증진하는 데 중점을 둔 정책을 마련할 여지를 제공할 수 있다. 가정, 사회 및 건강 설문조사에서 정신건강 문제 및 긍정적 정신건강에 대한 데이터를 수집하면 정신건강에 대한 더욱 완전한 이해를 도모할 수 있으며, 정신건강을 개선하는데 필요한 동인 및 정책 수단을 정확히 파악하는 데 도움이 될 것이다.

참고문헌

- American Psychiatric Association (2022), *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR™)*, <https://www.appi.org/dsm> (accessed on 6 April 2022). [64]
- American Psychiatric Association (2013), *Diagnostic statistical manual of mental disorder (5th ed.)*. [42]
- Australian Early Development Census (2012), *The mental health of Australian children: a dual continuum*, <https://www.aedc.gov.au/resources/detail/the-mental-health-of-australian-children-a-dual-continuum> (accessed on 17 May 2022). [20]
- Barlow, D. and V. Durand (2009), *Abnormal Psychology: An Integrative Approach*, Cengage Learning, https://books.google.com/books/about/Abnormal_Psychology_An_Integrated_Approa.html?id=Mo_q4zFVNo4C (accessed on 7 February 2022). [48]
- Conversano, C. et al. (2010), “Optimism and Its Impact on Mental and Physical Well-Being”, *Clinical Practice & Epidemiology in Mental Health*, Vol. 6/1, pp. 25-29, <https://doi.org/10.2174/1745017901006010025>. [56]
- Davydov, D. et al. (2010), “Resilience and mental health”, *Clinical Psychology Review*, Vol. 30/5, pp. 479-495, <https://doi.org/10.1016/j.cpr.2010.03.003>. [53]
- Exton, C. and L. Fleischer (2023), *The future of the How's Life? dashboard*, OECD Publishing. [61]
- Finnish Institute for Health and Welfare (n.d.), *Positive mental health*, <https://thl.fi/en/web/mental-health/mental-health-promotion/positive-mental-health> (accessed on 7 April 2022). [63]
- Fusar-Poli, P. et al. (2020), “What is good mental health? A scoping review”, *European Neuropsychopharmacology*, Vol. 31, pp. 33-46, <https://doi.org/10.1016/j.euroneuro.2019.12.105>. [59]
- Fusar-Poli, P. and Z. Santini (2022), “Promoting good mental health in the whole population: The new frontier”, *European Neuropsychopharmacology*, Vol. 55, pp. 8-10, <https://doi.org/10.1016/j.euroneuro.2021.09.010>. [58]
- Galderisi, S. et al. (2015), “Toward a new definition of mental health”, *World Psychiatry*, Vol. 14/2, pp. 231-233, <https://doi.org/10.1002/wps.20231>. [28]
- Goodman, F., J. Doorley and T. Kashdan (2018), “Well-being and Psychopathology: A Deep Exploration into Positive Emotions, Meaning and Purpose in Life, and Social Relationships”, in Diener, E., S. Oishi and L. Tay (eds.), *Handbook of well-being*, DEF Publishers, Salt Lake City. [30]
- Government of Canada (n.d.), *Positive Mental Health Surveillance Indicator Framework (PMHSIF)*, <https://health-infobase.canada.ca/positive-mental-health/Index> (accessed on 8 September 2022). [62]
- Government of Western Australia Mental Health Commission (2021), *Western Australian Mental Wellbeing Framework*, <https://www.mhc.wa.gov.au/mental-wellbeing-framework/> (accessed on 14 March 2022). [22]

- Greenblat, L. (2000), “Understanding Health as a Continuum”, *Schizophrenia Bulletin*, Vol. 26/1, pp. 243-245, <https://doi.org/10.1093/oxfordjournals.schbul.a033444>. [17]
- Harvard Center for Health and Happiness (n.d.), *Repository of Positive Psychological Well-Being Scales*, <https://www.hsph.harvard.edu/health-happiness/repository-of-positive-psychological-well-being-scales/> (accessed on 7 April 2022). [60]
- Herron, S. and D. Trent (2000), “Mental Health: A Secondary Concept to Mental Illness”, *Journal of Public Mental Health*, Vol. 2/2, pp. 29-38, <https://doi.org/10.1108/17465729200000014>. [14]
- Horwitz, A. (2007), “Distinguishing distress from disorder as psychological outcomes of stressful social arrangements”, *Health: An Interdisciplinary Journal for the Social Study of Health, Illness and Medicine*, Vol. 11/3, pp. 273-289, <https://doi.org/10.1177/1363459307077541>. [44]
- Huppert, F. (2005), “Positive mental health in individuals and populations”, in *The Science of Well-Being*, Oxford University Press, <https://doi.org/10.1093/acprof:oso/9780198567523.003.0012>. [49]
- Iasiello, M., J. van Agteren and E. Cochrane (2020), “Mental Health and/or Mental Illness: A Scoping Review of the Evidence and Implications of the Dual-Continua Model of Mental Health”, *Evidence Base*, Vol. 2020/1, pp. 1-45, <https://doi.org/10.21307/eb-2020-001>. [10]
- Keyes, C. (2005), “Mental Illness and/or Mental Health? Investigating Axioms of the Complete State Model of Health.”, *Journal of Consulting and Clinical Psychology*, Vol. 73/3, pp. 539-548, <https://doi.org/10.1037/0022-006X.73.3.539>. [11]
- Keyes, C., S. Dhingra and E. Simoes (2010), “Change in Level of Positive Mental Health as a Predictor of Future Risk of Mental Illness”, *American Journal of Public Health*, Vol. 100/12, pp. 2366-2371, <https://doi.org/10.2105/AJPH.2010.192245>. [25]
- Koushede, V. and R. Donovan (2022), “Applying Salutogenesis in Community-Wide Mental Health Promotion”, in *The Handbook of Salutogenesis*, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-030-79515-3_44. [18]
- New Zealand Initial Mental Health and Wellbeing Commission (2020), *Developing a Mental Health and Wellbeing Outcomes Framework Summary of what we heard through the co-define phase*, <https://www.mhwc.govt.nz/assets/Outcomes-framework/Co-define-Report-on-Responses-Online-Version.pdf> (accessed on 1 June 2022). [29]
- Nielsen, L. et al. (2016), “High and low levels of positive mental health: are there socioeconomic differences among adolescents?”, *Journal of Public Mental Health*, Vol. 15, pp. 37-49. [32]
- OECD (2021), *A New Benchmark for Mental Health Systems: Tackling the Social and Economic Costs of Mental Ill-Health*, OECD Health Policy Studies, OECD Publishing, Paris, <https://doi.org/10.1787/4ed890f6-en>. [3]
- OECD (2021), *COVID-19 and Well-being: Life in the Pandemic*, OECD Publishing, Paris, <https://doi.org/10.1787/1e1ecb53-en>. [5]

- OECD (2021), *Fitter Minds, Fitter Jobs: From Awareness to Change in Integrated Mental Health, Skills and Work Policies*, Mental Health and Work, OECD Publishing, Paris, [7]
<https://doi.org/10.1787/a0815d0f-en>.
- OECD (2020), *How's Life? 2020: Measuring Well-being*, OECD Publishing, Paris, [4]
<https://doi.org/10.1787/9870c393-en>.
- OECD (2019), *Health at a Glance 2019: OECD Indicators*, OECD Publishing, Paris, [40]
<https://doi.org/10.1787/4dd50c09-en>.
- OECD (2015), *2015 OECD Recommendation on Integrated Mental Health, Skills and Work Policy*, [6]
https://www.oecd.org/els/emp/Flyer_MHW%20Council%20Recommendation.pdf (accessed on 13 March 2022).
- OECD (2013), *OECD Guidelines on Measuring Subjective Well-being*, OECD Publishing, Paris, [52]
<https://doi.org/10.1787/9789264191655-en>.
- OECD (2011), *How's Life?: Measuring Well-being*, OECD Publishing, Paris, [8]
<https://doi.org/10.1787/9789264121164-en>.
- OECD/European Union (2018), *Health at a Glance: Europe 2018: State of Health in the EU Cycle*, [41]
 OECD Publishing, Paris/European Union, Brussels, https://doi.org/10.1787/health_glance_eur-2018-en.
- Patel, V. et al. (2018), “The Lancet Commission on global mental health and sustainable development”, [15]
The Lancet, Vol. 392/10157, pp. 1553-1598, [https://doi.org/10.1016/S0140-6736\(18\)31612-X](https://doi.org/10.1016/S0140-6736(18)31612-X).
- Payton, A. (2009), “Mental Health, Mental Illness, and Psychological Distress: Same Continuum or [16]
 Distinct Phenomena?”, *Journal of Health and Social Behavior*, Vol. 50/2, pp. 213-227,
<https://doi.org/10.1177/002214650905000207>.
- Peterson, C. and M. Seligman (2004), *Character Strengths and Virtues: A Handbook and Classification*, [55]
 American Psychological Association.
- Phillips, M. (2009), “Is distress a symptom of mental disorders, a marker of impairment, both or [45]
 neither?”, *World psychiatry : official journal of the World Psychiatric Association (WPA)*, Vol. 8/2,
 pp. 91-2.
- Queensland Government (2015), *The dual continuum of mental health and mental illness*, [23]
<https://cspm.csyw.qld.gov.au/practice-kits/mental-health/overview-of-mental-health-1/the-dual-continuum-of-mental-health-and-mental-ill> (accessed on 17 May 2022).
- Rehm, J. and K. Shield (2019), “Global Burden of Disease and the Impact of Mental and Addictive [2]
 Disorders”, *Current Psychiatry Reports*, Vol. 21/2, p. 10, <https://doi.org/10.1007/s11920-019-0997-0>.
- Reis, H. and S. Gable (2003), “Toward a positive psychology of relationships.”, in *Flourishing: Positive [51]
 psychology and the life well-lived.*, American Psychological Association, Washington,
<https://doi.org/10.1037/10594-006>.

- Robinson, P. (2012), “Deep phenotyping for precision medicine”, *Human Mutation*, Vol. 33/5, pp. 777-780, <https://doi.org/10.1002/humu.22080>. [26]
- Roger T. Mulder (2008), “An Epidemic of Depression or the Medicalization of Distress?”, *Perspectives in Biology and Medicine*, Vol. 51/2, pp. 238-250, <https://doi.org/10.1353/pbm.0.0009>. [46]
- Routledge, K. et al. (2016), “Shared versus distinct genetic contributions of mental wellbeing with depression and anxiety symptoms in healthy twins”, *Psychiatry Research*, Vol. 244, pp. 65- 70, <https://doi.org/10.1016/j.psychres.2016.07.016>. [12]
- Ryff, C. and C. Keyes (1995), “The structure of psychological well-being revisited.”, *Journal of Personality and Social Psychology*, Vol. 69/4, pp. 719-727, <https://doi.org/10.1037/0022-3514.69.4.719>. [57]
- Santini, Z. et al. (2022), “Higher levels of mental wellbeing predict lower risk of common mental disorders in the Danish general population”, *Mental Health & Prevention*, Vol. 26, p. 200233, <https://doi.org/10.1016/j.mhp.2022.200233>. [27]
- Santini, Z. et al. (2017), “The protective properties of Act-Belong-Commit indicators against incident depression, anxiety, and cognitive impairment among older Irish adults: Findings from a prospective community-based study”, *Experimental Gerontology*, Vol. 91, pp. 79-87, <https://doi.org/10.1016/j.exger.2017.02.074>. [37]
- Santini, Z. et al. (2020), “Predictors of high and low mental well-being and common mental disorders: Findings from a Danish population-based study”, *European Journal of Public Health*, Vol. 30/3, pp. 503-509, <https://doi.org/10.1093/eurpub/ckaa021>. [33]
- Snow, N. (2019), “Positive psychology, the classification of character strengths and virtues, and issues of measurement”, *The Journal of Positive Psychology*, Vol. 14/1, pp. 20-31, <https://doi.org/10.1080/17439760.2018.1528376>. [54]
- Solin, P. et al. (2019), “Predicting high and low levels of positive mental health among people living alone in the Lapland region of Finland”, *Journal of Public Mental Health*, Vol. 18, pp. 214-223. [38]
- Statistics Canada (n.d.), *Positive mental health and mental illness*, <https://www150.statcan.gc.ca/n1/pub/82-003-x/2014009/article/14086-eng.htm> (accessed on 17 May 2022). [19]
- Steger, M. et al. (2006), “The meaning in life questionnaire: Assessing the presence of and search for meaning in life.”, *Journal of Counseling Psychology*, Vol. 53/1, pp. 80-93, <https://doi.org/10.1037/0022-0167.53.1.80>. [50]
- Stewart-Brown, S. et al. (2015), “Socioeconomic gradients and mental health: implications for public health”, *British Journal of Psychiatry*, Vol. 206/6, pp. 461-465, <https://doi.org/10.1192/bjp.bp.114.147280>. [31]
- Swiss Health Observatory (n.d.), *Psychischer Gesundheitszustand [Mental Health Status]*, <https://www.obsan.admin.ch/de/gesundheits Themen/psychische-gesundheit/psychischer-gesundheitszustand> (accessed on 17 May 2022). [21]

- Thoits, P. (2011), “Mechanisms Linking Social Ties and Support to Physical and Mental Health”, [39]
Journal of Health and Social Behavior, Vol. 52/2, pp. 145-161,
<https://doi.org/10.1177/0022146510395592>.
- Trent, D. (1992), “The promotion of mental health: fallacies of current thinking”, in Trent, D. and C. [13]
 Reeds (eds.), *Promotion of Mental Health*, Aldershot, Avebury.
- Van Lente, E. et al. (2012), “Measuring population mental health and social well-being”, *International* [36]
Journal of Public Health, Vol. 57/2, pp. 421-430, <https://doi.org/10.1007/s00038-011-0317-x>.
- Varin, M. et al. (2020), “At-a-glance - An update on positive mental health among adults in Canada”, [34]
Health Promotion and Chronic Disease Prevention in Canada, Vol. 40/3, pp. 86-91,
<https://doi.org/10.24095/hpcdp.40.3.04>.
- Vinko, M. et al. (2022), “Positive mental health in Slovenia before and during the COVID-19 [35]
 pandemic”, *Frontiers in Public Health*, Vol. 10, <https://doi.org/10.3389/fpubh.2022.963545>.
- Wakefield, J. et al. (2007), “Extending the Bereavement Exclusion for Major Depression to Other [47]
 Losses”, *Archives of General Psychiatry*, Vol. 64/4, p. 433, <https://doi.org/10.1001/archpsyc.64.4.433>.
- World Health Organization (2022), *World mental health report: Transforming mental health for all*, [24]
<https://www.who.int/publications/i/item/9789240049338> (accessed on 7 December 2022).
- World Health Organization (2021), *International Classification of Diseases - ICD -11 for Mortality and* [43]
Morbidity Statistics.
- World Health Organization (2019), *Mental health*, [https://www.who.int/news-room/facts-in-](https://www.who.int/news-room/facts-in-pictures/detail/mental-health) [1]
[pictures/detail/mental-health](https://www.who.int/news-room/facts-in-pictures/detail/mental-health) (accessed on 16 May 2022).
- World Health Organization (1948), *WHO Constitution Preamble*, [9]
<https://www.who.int/about/governance/constitution> (accessed on 5 April 2022).

주

¹ 본 보고서에서 정신질환은 후술하는 바와 같이 진단 가능한 정신적, 행동적 상태는 물론 심리적 고통이라는 초진단적 특성을 의미하기도 한다. 본 보고서에서 정신건강 상태와 정신장애라는 용어는 임상적으로 심각한 증상을 동반하는 정신질환을 지칭하는 용어로 혼용하여 사용한다. 긍정적 정신건강은 심리적, 정서적, 경우에 따라 사회적, 관계적, 영적 웰빙을 포괄한다. 본 보고서는 주관적 웰빙의 관련 개념과 중첩되는 부분이 많고 2013년에 발간된 *OECD 주관적 웰빙 측정 지침*에서 심도 있게 다루었던 긍정적 정신건강 영역에 주로 초점을 맞추어 작성되었다. 해당 지침에서 긍정적 정신건강의 개념은 ‘사람들이 자신의 삶에 대해 긍정적 또는 부정적인 다양한 평가를 내리고 자신의 경험에 대해 정서적 반응을 보이는 양호한 정신 상태’로 정의하고 있다(OECD, 2013_[52]).

² 예를 들어 *How's Life? 2020*에서 우울증 위험군의 비율은 유럽건강면접조사(5년~6년마다 실시)에 조사 대상으로 포함된 유럽 국가에 한하여 보고되었으며, 부정적 정서 균형(긍정적 감정과 상태보다 부정적으로 보고되는 인구 비율)에 대한 정보는 현재 Gallup World Poll에서 제공된다. 마찬가지로 2021년 *새로운 정신건강 시스템 벤치마크 및 더 건강한 마음, 더 좋은 일자리* 연구 보고서에 포함된 정신적 고통의 불평등을 분석하는 데 사용된 설문조사 중 일부는 2015년 이전에 실시되었으며 다양한(통일되지 않은) 도구를 사용하여 고통을 측정한다.

³ 현재 웰빙은 11가지 영역으로 구성되며, 이는 사람들의 경제적 선택에 영향을 미치는 물질적 조건(소득과 부, 주거, 일과 일자리의 질)과 건강한 정도(그리고 건강하다고 느끼는 정도), 건강에 대한 지식과 역량, 삶의 터전이 건강하고 안전한 정도(건강, 지식과 역량, 환경의 질, 주관적 웰빙, 안전)를 포괄하는 삶의 질 요소와 관련되어 있다. 삶의 질은 또한 사회와 연결되어 있고 참여하는 정도, 그리고 시간을 보내는 방법(일과 삶의 균형, 사회적 관계, 시민 참여)을 포함한다. 미래의 웰빙을 위한 자원은 시간이 지나도 지속되지만 국가의 정책이나 조치와 관계없이 영향을 받는 다양한 유형의 자본 자원에 대한 국가의 투자(또는 고갈)로 표현된다. 여기에는 경제 자본(부 및 금융 자산), 자연 자본(천연 자원 재고, 토지 피복, 생물종의 다양성, 생태계 및 생태계 서비스), 인적 자본(개인의 역량 및 미래 건강) 및 사회 자본(사회 규범, 협력을 촉진하는 공유 가치 및 제도적 장치)이 포함된다(OECD, 2020_[4]).

⁴ OECD 웰빙 프레임워크와 함께 제공되는 지표 대시보드는 현재의 웰빙과 이를 유지하는 데 필요한 자원을 구분하여 두 영역에 대해 서로 다른 지표를 사용하므로 인적 자본 지표 세트에는 조기사망률과 비만유병률만 포함된다(Exton 및 Fleischer, 2023_[61]). 그러나 다른 영역에 포함되어 있는 사람들의 신체적, 정신적 건강은 개념적으로 인적 자본의 범위 내에 있고 이후의 삶의 기회에 영향을 미친다.

⁵ 이러한 관점은 WHO가 건강을 ‘완전한 신체적, 정신적, 사회적 웰빙 상태이고 단지 질병이 없는 상태 이상의 것’으로 광범위하게 정의한 데에도 반영되어 있다.

⁶ 이중 연속체 모델, 이중 요인, 이중 연속체, 완전 상태 모델, 완전 정신건강 등 이중 연속체 모델에 대한 다양한 명칭이 제안되었다.

⁷ 여기에서 정신건강의 유병률 또는 변영감은 쾌감 증상과 긍정적 기능의 증상으로 정의되었으며, 이는 긍정적 정서, Ryff의 심리적 웰빙 척도 및 Keyes의 사회적 웰빙 척도에 대한 6가지 문항으로 측정된다(Keyes, 2005_[11]).

⁸ 이러한 연구는 서구권 및 비서구권의 임상 및 비임상 모집단을 대상으로 전체 생애 과정에 대해 수행되었으며, 특히 소수 집단과 위험군을 대상으로 한 연구도 포함되었다.

⁹ 미국정신의학협회의 정신질환 진단 및 통계 편람(DSM)과 WHO 국제질병분류(ICD)에는 특정건강 상태를 진단하는데 필요한 일련의 기준이 명시되어 있다(세계보건기구, 2021_[43], American Psychiatric Association, 2013_[42]). 이러한 진단기준은 장애 유형에 따라 달라지며 증상의 성격과 수, 고통이나 손상의 수준을 지정하고 신체적 상해와 같은 일반적인 의학적 상태 또는 사랑하는 사람의 죽음과 같은 일반적인 스트레스 요인이나 상실에 대한 예측 가능하거나 문화적으로 수용될 수 있는 반응이 증상의 직접적인 원인인 경우는 제외하는 데 사용된다. DSM은 미국 정신의학 전문가들이 개발하여 주로 미국에서 사용되며 정신장애에 국한된 특정 분류 체계인 반면 ICD는 신체 장애와 정신장애를 모두 포괄하는 통합 분류 체계이다. DSM의 초판은 1952년에 발행되었으며 106개의 특정 진단을 포함했다. 이후 여러 차례 개정되어 2013년에는 총 157개의 진단과 300여 개의 장애 목록을 수록한 최신 개정판(DSM-5)이 발행되었다. 2022년 3월에는 진단 기준 및 진단 코드의 업데이트, 지속적 애도장애(Prolonged Grief Disorder)의 정신건강 상태로 신규 편입, 인종주의와 차별이 정신장애에 미치는 영향에 대한 고려 사항 등이 포함된 제5판 텍스트 개정판(DSM-5-TR)이 발행되었다(American Psychiatric Association, 2022_[64]). ICD에는 정신장애를 다루는 장(F장)이 있으며, 정기적으로 개정되어 2019년에 11차 개정판이 발간되었다. 두 체계는 약간의 차이는 있으나 유사한 규칙과 가정을 기반으로 한다.

¹⁰ 보건계량연구소(The Institute for Health Metrics and Evaluation)가 조사한 정신건강 장애에 대한 질병 부담 추정치(burden of disease estimates)는 다양한 출처의 데이터와 특정 장애 또는 위험 요인의 유병률, 삶의 질과 조기 사망률에 미치는 상대적 피해에 관한 일련의 복잡한 가정을 기반으로 한다.

¹¹ 긍정적 정신건강이 더욱 정교하고 광범위하게 개념화되는 방식은 OECD 국가의 여러 정부기관이 개념을 조작화하는 방식에서 분명히 드러난다. 캐나다 긍정적 정신건강 감시지표 프레임워크(Canadian Positive Mental Health Surveillance Indicator Framework)는 긍정적 정신건강을 ‘삶을 즐기고 직면한 문제에 대처하는 능력을 향상하는 방식으로 느끼고 생각하고 행동할 수 있는 웰빙상태’로 정의한다(Government of Canada, n.d._[62]). 반면 핀란드 보건복지연구원(Finnish Institute for Health and Welfare)에서는 이를 ‘다양한 수준의 정서적(감정), 심리적(긍정적 행동), 사회적(타인 및 사회와의 관계) 신체적(신체적 건강), 영적(삶에 의미가 있다는 느낌) 웰빙’으로 설명한다(Finnish Institute for Health and Welfare, n.d._[63]).

2 인구 정신건강 측정: 측정 도구 및 기존 국가별 측정 관행

행정적 데이터부터 다양한 유형의 설문조사 문항에 이르기까지 인구 정신건강을 모니터링하는 데 사용할 수 있는 다양한 도구가 있다. 코로나19 팬데믹 기간에 많은 OECD 국가에서 정신건강 데이터를 새롭게, 또는 추가로 수집하기 시작했으나 공식 데이터 생산기관에서는 팬데믹이 도래하기 훨씬 전부터 이미 해당 데이터에 대한 수집 활동을 활발히 실시해왔다. 그러나 (설문조사) 데이터 수집 빈도를 높이고 정신건강의 전체 스펙트럼을 포괄하기 위해 사용되는 지표 유형을 다양화 하며, 기존 척도의 국제적 통일화를 확대하는 등 개선이 필요하다. 여기서 데이터 수집기관은 다음과 같은 노력을 기울일 수 있다. (1) 우울증 증상에 초점을 맞춘 선별 도구를 넘어 불안 증상을 포함한 결과 측정 도구로 사용을 확대하고, (2) 긍정적 정신건강의 정서적 측면과 행복 측면에 대한 통일된 정보를 수집하는 방향으로 나아가며, (3) 설문조사 전반에 걸쳐 일반적인 정신건강 상태에 대한 단일 항목 문항의 사용을 모색할 수 있다.

인구 수준의 정신건강 결과 데이터를 자주 수집하는 것은 정신건강 위험군을 파악하고, 사람들의 정신건강을 형성하고 영향을 미치는 사회경제적 요인이나 기타 요인을 파악하며, 효과적인 예방 및 증진 전략을 설계하는 데 중요하다. 제1장에서 설명한 바와 같이 정신건강은 다면적인 개념이며, 정신질환의 유무라는 이분법적 구분을 넘어선 개념이다. 인구 조사 및 정신건강 평가에서 정신건강 문제 및 긍정적 정신건강에 대한 데이터를 수집하면 사람들의 전반적인 정신건강에 대해 더욱 완전한 이해를 도모할 수 있고, 정신건강의 개선과 관련된 동인 및 정책 수단을 정확히 파악하는 데 도움이 될 것이다.

그러나 현재 인구 정신건강에 대한 (국제적으로) 표준화된 데이터가 부족하기 때문에 통일성 없는 데이터를 토대로 다양한 정책 접근법의 효과를 평가하기 어려운 실정이다. 따라서 이러한 분석을 용이하게 하려면 결과 측정을 표준화하는 것이 첫 번째 단계이다. 본 장에서는 데이터 수집기관에서 사용할 수 있는 측정 도구를 개괄적으로 설명하고 OECD 국가 전반의 기존 데이터 수집 관행에 대한 개요를 제공하며, 국제적으로 통일된 데이터를 수집하기 위해 우선순위로 두어야 할 결과가 무엇인지 제안한다.

2022년 OECD 국가의 공식 데이터 생산기관에 전달된 설문지의 응답을 분석한 결과, 설문지에 응답한 모든 회원국에서는 이미 정신건강 데이터에 대한 수집 활동을 활발히 실시하고 있는 것으로 나타났다. 팬데믹 이전에도 거의 모든 OECD 회원국에서는 건강에 대한 면담 조사, 일반 가구 조사 및 행정적 데이터를 통해 정신건강 결과에 대한 정보를 수집해 왔다. 코로나19로 인해 인구 수준의 정신건강 측정에 대한 관심이 촉발되면서 많은 공공기관과 통계청에서 신규 또는 기존 설문조사에 정신건강에 관한 항목을 추가하였다.

이러한 데이터 수집은 국가 통계 차원에서 인구 정신건강 결과에 대한 관심을 가지고 있고 그것이 적합하다는 것을 보여준다. 그러나 데이터 수집 빈도, 부정적 및 긍정적 정신건강 결과의 전체 스펙트럼에 대한 데이터 가용성 향상, 국제적으로 비교가능한 국가 간 측정의 통일화 등 몇 가지 영역에서 개선의 여지가 있다.

실제로 팬데믹 이전에는 대부분의 정신건강 데이터가 4년~10년마다 실시되는 설문조사를 통해 국가별로 수집되었다. 코로나19의 출현 이후 초기 2년 동안 많은 국가에서 정신건강 모듈이 포함된 설문조사를 도입하여 자주 조사를 실시하고 있으나, 이러한 고빈도 설문조사가 앞으로도 지속될지는 현재로서 불분명하다. 또한 모든 통계청에서 일반적인 정신장애에 중점을 두고 정신건강 문제에 대한 데이터를 수집하고 있지만 일반적인 심리적 고통과 우울 증상은 표준화된 선별 도구를 통해 파악하는 경향이 있는 반면, 불안 경험에 대한 측정은 국가 간에 통일성이 결여되어 있는 실정이다. 외상 후 스트레스 장애(PTSD), 양극성 장애, 섭식장애 등 기타 정신질환과 자살 충동 및 정신건강 관련 낙인과 같은 정신건강의 다른 측면에 대한 데이터의 수집은 국가별로 균등하지 않은 상태이다. 긍정적 정신건강의 경우 국가 간 비교 데이터는 주로 생애 평가 척도에 국한되어 있다. 정서 및 행복 같은 측면은 결과 측정으로 수집되는 빈도가 매우 낮으며, 수집되는 경우에도 사용되는 측정 도구가 국가 간에 표준화되어 있지 않다.

OECD 설문지의 결과에 따르면 기존의 데이터 수집 관행이 긍정적 정신건강뿐만 아니라 정신건강 문제의 측면도 누락되어 전체 범위의 정신건강 결과를 포착하지 못하고 있는 것으로 나타났다. 전반적인 정신건강 결과를 포착하고 정신건강에 대한 정보 수집의 빈도를 높이기 위해 OECD 회원국의 데이터 수집기관은 다음과 같은 노력을 기울일 수 있다. (1) 우울증 증상에 초점을 맞춘 선별 도구를 넘어 불안 증상을 포함한 결과 측정 도구로 사용을 확대하고, (2) 긍정적 정신건강의 정서적 측면과 행복 측면에 대한 통일된 정보를 수집하는 방향으로 나아가며, (3) 설문조사 전반에 걸쳐 일반적인 정신건강 상태에 대한 단일 항목 문항의 사용을 모색할 수 있다.

인구 정신건강을 측정하는 데 사용할 수 있는 도구는 무엇인가?

제1장에서는 데이터 수집기관이 정신건강에서 관심을 가져야 하는 결과의 유형에(정신건강 문제 및 긍정적 정신건강 모두 포함)에 중점을 두었다면 본 장에서는 해당 결과를 측정하는 데 사용할 수 있는 도구 유형에 대해 다룬다.

본 장에서 논의하는 광범위한 도구 유형 중 일부는 행정적 데이터에서 수집되었으며, 대부분의 경우 가구 조사에서 수집된 것으로 설문조사는 확장형 모듈부터 문항 항목 모음, 단일 문항에 이르기까지 다양하다. 일부 도구는 정신건강 문제 또는 긍정적 정신건강의 측면을 포착하는 데 사용할 수 있는 반면, 특정 유형의 결과에 한하여 사용되는 도구도 있다. 각 도구 유형마다 고유한 장점과 단점이 있으므로 데이터 수집기관은 특정 상황에서의 요구사항과 제약 조건에 따라 도구를 선택해야 한다. 본 보고서에 사용된 분류에 대한 공통적인 이해를 돕기 위해 다양한 유형의 도구가 아래에 설명되어 있다.

각 장의 부속서에는 보다 자세한 내용을 독자들에게 제공하기 위해서 심층 정보가 포함되어 있다. 부속서 2.A에서는 국가별로 수집된 특정 도구에 대한 개요와 함께 문항 구성 및 답변 옵션 샘플을 제공한다. 부속서 2.B에서는 가장 일반적으로 사용되는 표준화된 도구에 대한 문항 문구 및 채점 권장사항을 포함한 전체 세부 정보를 제공한다. 정신건강 설문조사의 통계 품질에 대한 자세한 내용은 제3장에서 다룬다.

행정적 데이터에서 수집된 도구

행정적 데이터에는 **정신건강 서비스 이용, 임상 환경에서의 정신장애 진단, 자살 및 약물 남용으로 인한 사망 원인 데이터**(예: 약물 과다복용 및 알코올 남용)에 대한 정보가 포함될 수 있다.

이러한 데이터는 객관적이고(즉 자가 보고가 아님) 수집하기 용이한 정신건강 측정 지표로 간주될 수 있으나 해당 측정에는 개선할 여지가 남아있다. 예를 들어 의료 서비스 이용 및 의학적 진단에 대한 측정은 인구 수준의 결과를 포착하는 것이 아니라 의료 서비스에 이용할 의향이 있고 서비스를 이용할 수 있는 대상만을 포착한다. 이러한 측정은 우수하고 저렴한 의료 시스템과 인식개선 프로그램을 갖추고 있고 정신질환에 대한 대중의 낙인이 적어 정신질환 치료를 찾거나 받을 가능성이 높은 국가의 비교 수준이나 발병률을 과대평가할 수 있다. 또한 질환을 예방하려면 의료 서비스를 이용하기 전과 후의 결과를 추적해야 한다. 본 보고서에서는 보건의료와 관련된 행정적 통계는 더 이상 고려하지 않는다(OECD, 2021^[1] 참조)

자살 또는 약물 남용으로 인한 사망 원인에 대한 데이터[(일반적으로 ‘절망사’라고 칭함(Case 및 Deaton, 2017^[2]))]는 인구 수준에서 정신질환 결과를 파악할 수 있다. 이러한 척도는 심각한 정신질환 및 중독에 대한 측정 지표로 작용할 수 있다. 자살 행동에 영향을 미치는 사회적, 문화적 이유가 있으므로 - 모든 자살의 직접적인 원인이 정신질환은 아니지만 - 정신질환을 앓고 있는 경우 자살로 사망할 위험이 현저히 높아진다(OECD, 2021^[1]). 그러나 자살 사망 신고는 사망 원인의 확인, 사망진단서 작성자, 자살에 대한 통념 및 낙인 등 요인의 영향을 받는 복잡한 절차이며, 이는 사망 기록의 국가 간 비교 가능성에 잠재적으로 영향을 미칠 수 있다(OECD, 2021^[1]).

일반적으로 모든 유형의 행정적 데이터는 추가로 함께 수집되는 사회인구통계학적 데이터가 연령, 성별, 지리적 영역 및 잠재적으로 인종 또는 민족으로 제한되는 경우가 많다는 한계가 있다. 이로 인해 정신건강의 동인을 탐구하고 관련 사회경제적, 환경적, 관계적 위험 및 회복력 요인을 파악하는 데 제약이 있다.

가구 조사에서 수집된 도구

행정적 데이터와는 달리 인구 조사는 일반적으로 응답자의 물질적 조건(예: 소득, 부, 노동 시장 성과, 주거의 질), 삶의 질(예: 신체건강, 교육 수준, 환경의 질) 및 관계(예: 사회적 관계, 신뢰, 안전)에 대한 정보를

포함한다. 인구 조사는 건강 조사와 같은 특정한 내용에 초점을 맞출 수 있고 종합 사회 조사와 같은 보다 일반적인 범위를 가질 수도 있다. 이러한 조사는 고용, 건강(정신건강 포함), 교육 등에 대한 보다 심층적인 모듈을 사용해 가구 수준에서 실시된다. 모든 범위의 웰빙 공변량을 확보하는 것은 정신건강에 영향을 미치는 요인과 궁극적으로 정신건강이 삶의 다른 영역에 미치는 영향에 대해 이해하는 데 중요하다. 또한 정신건강 결과의 형평성을 추적하고 궁극적으로 이를 달성하려면 중요한 사회인구통계학적 범주별로 세분화해야 한다.

특정 정신건강 결과를 평가하기 위해 가구 조사에 포함된 도구는 단일 항목 문항부터 표준화된 항목 모음까지 다양하다. 각 도구에 대한 간략한 설명은 아래에서 확인할 수 있으며, 자세한 내용은 부속서 2.A 및 부속서 2.B에서 확인할 수 있다.

- **이전 진단에 관한 문항** - 지난 12개월 동안 또는 일생에 걸쳐 의료진으로부터 정신건강 장애(예: 주요 우울장애, 범불안장애 또는 기타 정신건강 상태)로 진단받은 적이 있는지에 대한 단일 항목 문항을 말한다. 해당 문항은 일반적으로 응답 방식이 ‘예/아니요’로 진행되며 국가별로 표준화되어 있지 않다. 자세한 내용을 확인하려면 표 2.6을 참조한다. 예시는 다음과 같다.
 - “귀하는 의료 전문가(정신과 전문의, 의사, 임상심리사)로부터 정신건강 문제를 정신장애로 진단받은 적이 있습니까? 예/아니요”
 - “귀하는 의사 또는 기타 의료 전문가로부터 어떤 유형의 우울증을 진단받은 적이 있습니까? 필요한 경우 다음을 참고하세요. 일반적인 우울증 유형에는 주요우울증(또는 주요우울장애), 양극성우울증, 기분부전증, 산후우울증 및 계절성정동장애가 있습니다. 예/아니요”
- **경험한 증상에 관한 문항** - 의료 전문가의 진단을 명시적으로 언급하지 않고 지난 12개월 동안 또는 일생에 걸쳐 경험한 정신장애의 증상에 대한 단일 항목 문항을 말한다. 해당 문항은 일반적으로 응답 방식이 ‘예/아니요’로 진행되며 국가별로 표준화되어 있지 않다. 자세한 내용을 확인하려면 표 2.7을 참조한다. 예시는 다음과 같다.
 - “지난 12개월 동안 다음 질병이나 상태를 경험한 적이 있습니까? 우울증(예/아니요)”
 - “만성 불안을 경험한 적이 있습니까? (예/아니요)”
 - “기분 장애를 가지고 있습니까? (예/아니요)”
- **자살 사고 및 자살 시도에 관한 문항** - 응답자의 자살 사고, 자해 행동 또는 자살 시도 경험에 대한 (일반적으로) 단일 항목 문항이다. 해당 문항은 일반적으로 응답 방식이 ‘예/아니요’로 진행되며 국가별로 표준화되어 있지 않다. 회상 기간은 개인의 일생, 지난 12개월, 지난 2주 또는 ‘코로나19 팬데믹 기간’ 이 있다. 자세한 내용을 확인하려면 표 2.8을 참조한다. 예시는 다음과 같다.
 - “코로나19 팬데믹이 시작된 이후 심각하게 자살을 고려한 적이 있습니까? 예/아니요”
 - “때때로 사람들은 자살할 의도 없이 자기 신체 일부를 고의로 상하게 하는 행동을 합니다. 귀하는 지난 12개월 동안 자살할 의도 없이 자해 행동을 한 적이 있습니까? 예/아니요”
 - “자살을 시도한 적이 있습니까? 예/아니요”
 - “귀하는 자살을 시도하여 하룻밤 이상 병원에 입원한 적이 있습니까? 예/아니요”
- **일반적인 정신건강 상태에 관한 문항** - 응답자가 자신의 정신건강을 전반적으로 어떻게 평가하는지에 대한 단일 항목 문항으로 정신건강 문제와 긍정적 정신건강의 두 가지 요소를 모두 파악할 수 있다. 문항은 국가별로 표준화되어 있지 않으며 문항 문구, 응답 옵션 및 회상 기간이 다르다. 자세한 내용을 확인하려면 표 2.9를 참조한다. 예시는 다음과 같다.

- “일반적으로 귀하의 정신건강은 어떻습니까? 훌륭하다/매우 좋다/좋다/보통이다/나쁘다”
- “지난 12개월 동안 코로나19 팬데믹으로 인해 귀하의 정신건강/웰빙에 영향을 받은 적이 있습니까?”
- “귀하의 정신건강에 어느 정도 만족하는지 1점부터 10점까지의 척도로 표시할 수 있습니까? 1점은 전혀 불만족, 10점은 매우 만족을 의미합니다.”
- “귀하의 정신 상태가 직장 생활이나 가정 생활에 지장을 초래합니까? 예/아니요”
- **긍정적 정신건강 지표** - 생애 평가, 정서(정서 척도 요약 및 긍정, 부정 또는 혼합 정서에 관한 문항 모음), 행복(삶의 질, 삶이 가치 있고 의미 있는지에 관한 문항), 표준화된 긍정적 정신건강 복합 척도(긍정적 정신건강의 여러 영역을 결합하여 부정적 정서보다 긍정적 정서를 우선시하고 경우에 따라 사회적 웰빙 요소를 추가) 등 긍정적 정신건강의 다양한 측면과 관련된 문항을 말한다. 일부 경우에는 긍정적 정신건강 지표가 국가 및 설문조사에 따라 달라지는 단일 항목 문항으로 구성되는 반면, 표준화된 문항 모음으로 구성되는 경우도 있다. 국가별 표준화는 다양하며 생애 평가에 관한 문항과 긍정적 정신건강 복합 척도에 한해 일관된 문구로 구성되어 있다. 자세한 내용을 확인하려면 표 2.10을 참조한다. 표준화된 복합 척도에 대한 구체적인 문항 문구 및 채점 권장사항은 부속서 2.B에서 확인할 수 있다.
- **선별 도구** - 정신건강 상태의 증상(진단이 아닌)에 따라 응답자를 선별하기 위해 고안된 다항목 도구를 말한다. 해당 도구는 임상 환경에서 일반적인 정신장애를 선별하여 위험군을 식별하고 추가 선별 및 잠재적 진단을 표시하기 위해 처음 개발되었다. 일반적인 심리적 고통 또는 주요우울장애, 범불안장애(이 두 가지가 복합적으로 나타나는 경우도 있음), 알코올 사용장애, 외상 후 스트레스 장애, 섭식장애 등과 같은 특정 정신건강 상태에 초점을 맞춰 면담자 주도식이나 자기 보고식으로 실시할 수 있다. 이러한 도구는 타당도(구조화된 면담 또는 진단의 표준 기준에 대한), 민감도(해당 질환을 가진 환자를 정확하게 식별할 확률) 및 신뢰도(주어진 상황에서 개인 면담 시 측정의 결과가 일관성 있게 산출되는 정도)에 대해 심리측정적으로 검증되었다는 점에서 ‘타당한’ 것으로 간주된다(통계 품질에 대한 자세한 논의는 제3장 참조). 두 가지 항목으로 구성된 단축형 도구부터 20개 이상의 항목을 포함하는 확장형 도구에 이르기까지 다양한 선별 도구를 사용할 수 있다. 선별 도구마다 질문의 초점은 다르지만 모든 도구에서 (대부분 부정적인) 정서(즉 기분저하, 초조함, 무가치감)를 경험하는 빈도를 다루며, 일부 도구에서는 신체 증상(예: 식욕 변화, 수면장애) 및 정서적 고통으로 인한 기능 손상(예: 일상 활동에서의 장애, 집중이 불가능, 걱정이 지속됨)를 포함한다. 선별 도구는 증상의 기준 기간도 지난주에서 지난달 까지 다양하지만 평생유병률을 측정할 수 없다. 선별 도구 간에는 이러한 차이가 있으므로 항상 직접 비교할 수 있는 것은 아니며, 국제 비교를 위해 상호 교환적으로 사용해서는 안 된다. 항목 점수는 일반적으로 요약 지수로 요약되며 최종 점수는 정신건강 문제의 연속적인 척도로 사용되거나 검증된 준거점수를 사용하여 일반적인 정신건강 상태의 위험을 평가하는 데 사용된다. 자세한 내용을 확인하려면 표 2.5를 참조한다. 가장 자주 사용되는 선별 도구에 대한 정확한 문항 문구 및 채점 권장사항은 부속서 2.B에서 확인할 수 있다.
- **구조화된 면담** - 구조화된 면담은 정신장애를 측정하기 위한 표준 기준으로 간주된다(일생 및 12개월 기준으로 실시하는 경우가 많음). 구조화된 면담은 국제적으로 합의된 정신의학적 분류 체계의 정의와 기준에 따라 표준화된 평가를 제공하며, 응답자가 해당 질환을 가지고 있는지를 판단하는 강력한 진단 신뢰도와 심리측정적 성을 가지고 있다(Mueller 및 Segal, 2015^[3], Burger 및 Neeleman, 2007^[4]).¹ 훈련된 면담자가 미리 정해져 있는 폐쇄형 문항을 통해 진행하며, 응답에 대해 표준화된 점수표에 따라 채점한다(Ruedgers, 2001^[5]).

구조화된 면담은 정신건강 전문가가 실시하는 평가와 유사하므로 의료 전문가로부터 진단받지 않았더라도 정신건강 상태의 위험군을 식별할 수 있다. 가장 일반적으로 사용되는 구조화된 면담인 복합 국제진단면담도구(Composite International Diagnostic Interview, CIDI)에 대한 자세한 내용을 확인하려면 표 2.4 및 부속서 2.B를 참조한다.

- **기타 정신건강 관련 주제** 이 범주는 정신건강 약물 및 서비스 이용, 가정 내 아동 및 청소년의 정신건강, 외로움 및 스트레스, 회복력 및 자기 효능감, 정신건강에 대한 낙인과 이해력을 포함하는 태도, 충족되지 않은 요구사항에 대한 문항 등 기타 관련 주제에 대한 문항이 포함된다. 자세한 내용을 확인하려면 표 2.11을 참조한다.

도구 유형 간의 상충관계

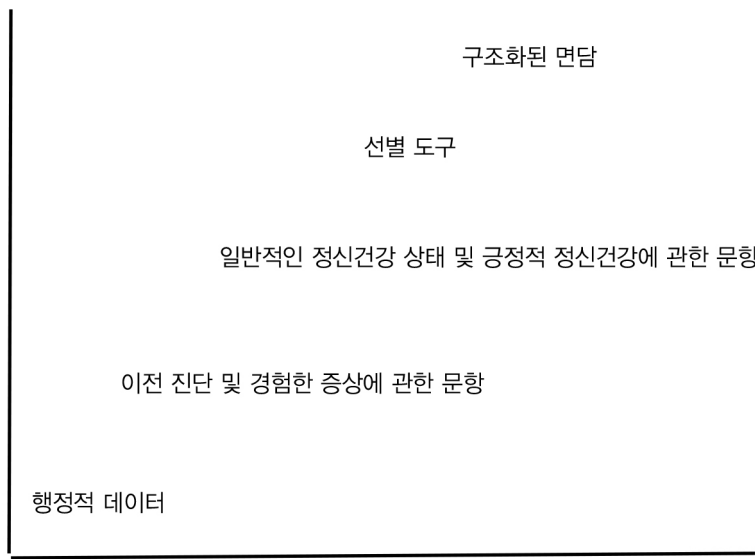
모든 도구는 응답 부담 또는 용이성 및 데이터 수집 비용, 정확도 및 적용범위 측면에서 상충관계를 내포하고 있다(그림 2.1, 표 2.1). 응답 부담은 설문조사 응답자가 자신의 정신건강 상태에 대한 정보를 제공하는 데 소비해야 하는 시간과 해당 정보를 제공함으로써 발생하는 스트레스와 직접적인 연관성이 있다. 정확도는 정신질환이 있는 사람을 정확하게 식별하는 도구의 민감도를 의미하며, 적용범위는 해당 도구가 전체(성인) 인구에 적용되는지를 의미한다.

예를 들어 행정적 데이터는 개인의 응답에 기초한 것이 아니라 국가의 데이터 인프라 내에서 일상적으로 수집되기 때문에 응답 부담이 적다. 그러나 절망사에 대한 통계는 정신질환의 극단적인 결과에만 초점을 맞추고 있으며, 모든 절망사가 정신장애의 정점을 나타내는 것은 아니므로 더욱 복잡하다. 또한 가구 조사와는 달리 의료 시스템을 이용한 대상만이 임상 환경에서 진단받은 것으로 행정 기록에 포착된다.²

가구 조사의 경우 도구의 문항 수가 많고 구체적일수록 응답 부담과 정확도가 모두 높아진다. 경험한 증상이나 개인의 일반적인 정신건강 상태에 대한 단일 문항은 수가 적고 대답하기 쉽지만, 증상의 본질이나 심각도, 정신건강 상태의 유형은 고려하지 않으며 진단 기준에 대해서도 검증하지 않는다. 선별 도구는 구조화된 면담의 표준 기준에 따라 타당화 되었으며, 특정 도구의 사용과 항목 수를 고려해 볼 때 응답 부담 측면에서 여전히 상대적으로 비용이 적게 든다. 그러나 이는 의료 전문가의 진단을 대신할 수 없으며 장애 위험군만 식별할 수 있을 뿐이다. 선별 도구는 임상 진단과 비교하여 타당화 되었으며, 따라서 가능한 진단 면담과 유사성을 극대화하여 설계되었다. 그럼에도 불구하고, 도구와 준거점수 보정 시 민감도(정신건강 상태가 있다는 것을 정확하게 식별)와 특이도(정신건강 상태가 없다는 것을 정확하게 표시) 사이에는 상충관계가 있다. 연구자들은 후자보다 전자를 우선시하는 경향이 있어 설계상 약간의 과대평가가 발생할 수 있다(자세한 내용은 상자 3.3 및 섹션 3.3.1 참조). 마지막으로, 가구 조사 및 행정적 데이터에 포함된 대부분의 도구는 정신질환에 초점을 맞추고 있다. 단, 일반적인 정신건강 및 긍정적 정신건강에 대한 가구 조사 문항은 예외이다.

경험한 증상이나 진단받은 질환에 대한 구조화된 면담, 선별 도구 및 단일 항목 문항 간의 문항 구성과 항목 길이의 차이로 인해 보고된 동일한 결과 측정에 대한 유병률 추정치가 달라질 수 있다(상자 2.1). 이는 국가별, 시간 경과별 및 인구 집단별로 결과를 비교 측정할 때 도구 유형의 표준화(및 사용된 도구에 대한 투명성)가 필요함을 시사한다. 즉 도구 유형을 혼합하여 ‘우울증 위험군 비율’ 또는 ‘심리적 고통의 위험 비율’과 같은 결과에 대해 언급하면 추정치가 근본적인 정신건강 상태의 차이가 아닌 측정 차이로 인해 달라질 수 있다(해당 주제에 대한 자세한 내용은 제3장 참조).

그림 2.1. 정신건강 측정 도구별 응답 부담과 정확도 간의 관계



정확도/적용범위

출처: 2021년 12월에 개최된 OECD 회의에서 캐나다 통계청이 발표한 프리젠테이션 “Well-being and mental health - towards an integrated policy approach”에서 인용

표 2.1. 정신건강 측정 도구별 장점 및 한계점

| 도구 | 장점 | 한계점 |
|---|--|--|
| 행정적 데이터(자살, 약물 과다 복용, 알코올 남용으로 인한 절망사, 임상 환경에서 일반적인 정신장애의 진단) | <ul style="list-style-type: none"> • 개인의 응답 부담 없음 • 다른 행정적 데이터(예: 의료 시스템 품질)와의 연계 가능성 • 다른 유형의 데이터보다 비용이 적게 들고 쉽게 이용할 수 있음 • 임상 진료 데이터는 다른 데이터 소스를 사용할 수 없을 경우 일생 및 특정 기간(예: 지난 12개월) 동안 발생한 다양한 질환의 유병률 추정치에 대한 통찰력을 제공할 수 있음 | <ul style="list-style-type: none"> • 의료 전문가에 의해 정확하게 코딩되고 데이터베이스에 기록된 일부, 치료받은 사람들에 대한 정보만 포착됨 • ‘사망 원인’ 데이터는 올바르게 코딩되어야 하며, 자살 시도나 약물 남용이 사망으로 이어지지 않은 경우를 고려하지 않아야 하고, 정신질환이 극단적 상황으로 이어진 경우만을 포착해야 함 • 진단율의 변화가 근본적인 정신건강 상태의 유병률의 변화에 기인하는 것인지 또는 치료의 경제성 또는 접근성의 변화, 도움을 구하는 행동의 변화 등과 같은 다른 요인에 기인한 것인지 (추가 정보 없이) 해석하기 어렵거나 불가능한 경우가 많음 • 웰빙 공변량에 대한 상황 정보가 제한됨 |
| 가구 조사: 이전 진단에 관한 문항 | <ul style="list-style-type: none"> • 응답자가 비교적 이해하기 쉬움 • 최소한의 응답 부담(일반적으로 단일 이분 문항) • 일생 및 특정 기간(예: 지난 12개월) 동안 발생한 다양한 질환의 유병률 추정치를 제공할 수 있음 | <ul style="list-style-type: none"> • 의료 전문가에게 치료를 받았거나 진단을 받은 대상만을 포착함 • 이러한 문항이 사회적 바람직성 편향과 높은 응답거부 및 무응답률로 이어진다는 증거가 있음(제3장 참조) • 증상의 성격과 심각도에 대한 상황 정보가 제한됨 |
| 가구 조사: 경험한 정신건강 상태의 증상에 관한 문항 | <ul style="list-style-type: none"> • 최소한의 응답 부담(일반적으로 단일 이분 문항) • 일생 및 특정 기간(예: 지난 12개월) 동안 발생한 다양한 질환의 유병률 추정치를 제공할 수 있음 | <ul style="list-style-type: none"> • 이러한 유형의 도구가 의료 전문가의 이전 진단에 관한 문항과 밀접하게 관련되어 있다는 증거가 있지만, 문항이 실제 진단을 나타내는지 또는 자기 평가를 나타내는지 응답자가 혼동할 가능성이 있음 • 증상의 성격과 심각도에 대한 상황 정보가 제한됨 |
| 가구 조사: 일반적인 정신건강 상태에 관한 문항 | <ul style="list-style-type: none"> • 응답자가 비교적 이해하기 쉬움 • 최소한의 응답 부담(일반적으로 단일 문항) | <ul style="list-style-type: none"> • 완전한 평가나 실제 진단이 아닌 실제 유병률에 대한 과대 보고, 증상을 고려하지 않음 |

| 도구 | 장점 | 한계점 |
|---------------------------|--|---|
| | <ul style="list-style-type: none"> • 응답자의 정신 상태에 대한 전반적인 평가, 즉 정신건강 문제와 긍정적 정신건강 측면을 모두 포착함 | <ul style="list-style-type: none"> • 구조화된 면담 또는 기타 진단 도구와 비교하여 타당화되지 않았으며, 위험군을 구성하는 요소에 대한 임계값이 도구에 설정되어 있지 않음 • 일반적으로 기존 증거 기반이 부족하지만 여러 연구에서 유용한 척도로 간주함 • 증상의 성격과 심각도 또는 정신건강 상태 유형에 대한 상황 정보가 제한됨 |
| 가구 조사: 긍정적 정신건강 지표 | <ul style="list-style-type: none"> • 응답자가 비교적 이해하기 쉬움 • 최소한의 응답 부담(일반적으로 단일 또는 제한된 항목 문항) • 심리적, 정서적 웰빙 또는 번영감에 중점을 둠 • 국제 측정 지침이 존재함(예: <i>OECD 주관적 웰빙 측정 지침</i>) | <ul style="list-style-type: none"> • (실제 또는 목표) 유병률에 대한 기준점이 없음 • 문항의 회상 기간은 일반적으로 하루에서 지난 4주의 기간이며, 일생의 추정치를 제공할 수 없음 |
| 가구 조사: 선별 도구 | <ul style="list-style-type: none"> • 구조화된 면담에 비해 실시하기가 쉽고 응답 부담이 적음 • 구조화된 면담 또는 기타 진단 도구와 비교하여 타당화됨 • 진단되지 않은 상태도 포착 가능함 | <ul style="list-style-type: none"> • 완전한 평가나 실제 진단이 아닌 실제 유병률에 대해 과대 보고 • 문항의 회상 기간은 일반적으로 하루에서 지난 4주의 기간이며, 일생의 추정치를 제공할 수 없음 |
| 가구 조사: 구조화된 면담 | <ul style="list-style-type: none"> • 표준 기준에 가까운 실제 유병률 근사치 제공 • 진단되지 않은 상태도 포착 가능함 • 응답자의 삶에 대한 광범위한 상황 정보를 고려할 수 있음 | <ul style="list-style-type: none"> • 면담자 훈련을 포함하여 개발 및 실시가 매우 복잡함 • 유증상자에 대한 문항 비율이 높음 • 최신 진단 지침인 정신질환 진단 및 통계 편람 제5판(DSM-5)에 매핑할 수 있는 설문조사 측정 도구가 부족함 |

출처: 2021년 12월에 개최된 OECD 회의에서 캐나다 통계청이 발표한 프리젠테이션 “Well-being and mental health - towards an integrated policy approach”에서 인용

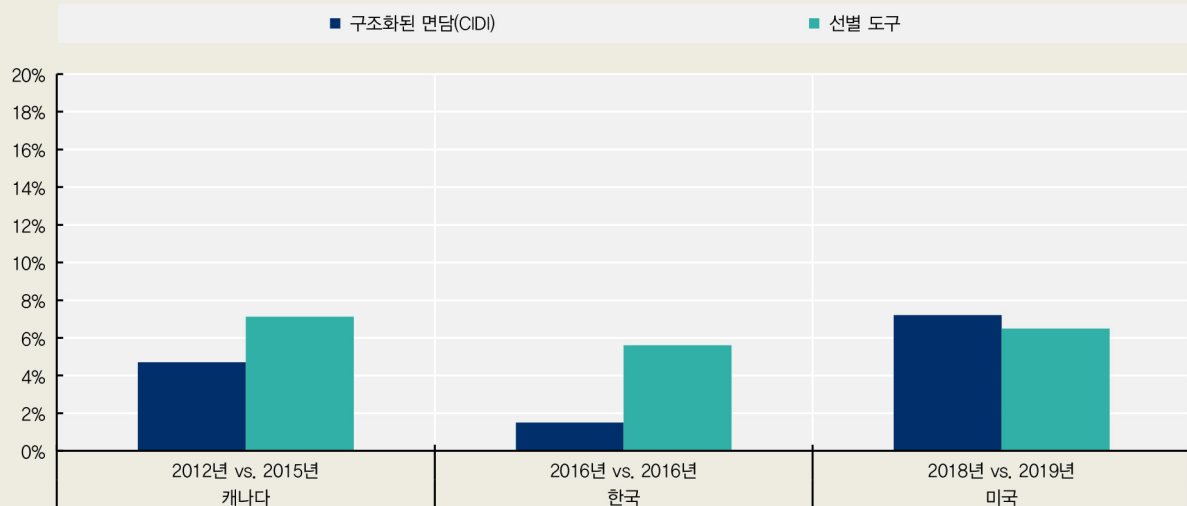
상자 2.1. 유병률은 사용하는 측정 도구에 따라 다르게 나타난다

특정 정신건강 상태의 유병률은 추정치를 산출하는 데 사용되는 도구의 유형에 따라 크게 달라질 수 있다. 선별 도구에 따라 설계상 인구 수준의 정신장애 유병률이 *과장될* 가능성이 높다. 선별 도구는 일반적인 정신장애의 위험군을 식별하기 위해 임상 환경에서 개발되었으며, 추가 관찰이 필요하거나 실제 진단받은 대상을 표시할 수 있다. 이들 중 일부는 진단되지 않거나 추가 치료가 필요하지 않을 수 있다(National Academies of Sciences Engineering and Medicine, 2021^[6], Topp 외, 2015^[7]). 반면 과거에 의료 전문가로부터 정신장애 진단을 받은 적이 있는지, 또는 현재 특정 장애를 겪고 있는지에 대한 문항은 의료 시스템을 이용하는 대상에 초점을 맞추기 때문에 인구 수준의 유병률을 *과소평가할* 가능성이 높다.

첫 번째로, 아래의 그림 2.2는 선별 도구가 구조화된 면담에 비해 인구 수준의 유병률을 과대평가할 수 있다는 점을 보여준다. 이 그림은 구조화된 면담인 복합 국제진단면담도구(CIDI)와 선별 도구인 환자 건강 설문지(PHQ)로 측정된 OECD 3개국의 주요우울장애(MDE) 유병률에 대한 국가별 측정 추정치를 보여준다. (PHQ는 국가별로 다른 버전을 사용하였으며 캐나다와 한국은 PHQ-9, 미국은 PHQ-2를 사용함, 구체적인 항목은 부속서 2.B 참조). 복합 국제진단면담도구(CIDI)와 선별 도구는 각 국가 내에서 서로 다른 설문조사에 사용되기 때문에 직접 비교하는 데 주의해야 하지만, 일반적으로 복합 국제진단면담도구(CIDI)로 측정된 주요우울장애(MDE) 유병률은 선별 도구로 측정된 유병률보다 낮다. 예외적으로 미국은 추정치 간의 차이가 가장 적은 것으로 나타났다. 이는 많은 정신건강 조사 도구가 미국에서 처음 개발되었고 이후 집중적인 타당화 과정을 거치며 서로 다른 도구 간에 보다 정확하게 보정되었기 때문이다.

그림 2.2. 일반적으로 선별 도구는 구조화된 면담에 비해 주요우울장애 유병률이 높게 나타난다

복합 국제진단면담도구(CIDI) 및 선별 도구(PHQ)로 추정된 지난 12개월 동안 및 지난 2주 동안의 주요우울삽화(MDE) 유병률



주: 세 국가 모두 구조화된 면담인 CIDI를 사용했으며, 이는 지난 12개월 동안의 주요우울삽화(MDE)의 유병률을 측정하는데 사용된다. 한국의 경우 해당 추정치가 연령과 성별에 맞게 조정되었다. 캐나다와 미국의 경우, 해당 추정치는 각각 15세 이상과 18세 이상 인구에 대해 국가적으로 대표성을 갖는다. 캐나다에서 사용하는 검증된 선별 도구는 PHQ-9(주요우울삽화가 10점 이상인 것으로 정의)이다. 한국에서는 PHQ-9(우울증 위험은 10점 이상인 것으로 정의되며, 한국의 KOSIS, 국가통계포털에서는 주요우울삽화 위험으로 정의하지 않았지만, 캐나다에서는 동일한 채점 기준을 적용하여 주요우울삽화를 측정함)를 사용하고 있다. 미국에서는 PHQ-2(우울장애 증상은 3점 이상인 것으로 정의됨)를 사용하고 있다. PHQ-9와 PHQ-2의 기준 기간은 모두 지난 2주이다. 미국의 경우 PHQ-2는 주요우울삽화의 경험보다는 우울장애 증상 유무의 비율을 측정한다. 개별 선별 도구에 대한 자세한 내용은 부속서 2.B를 참조한다.

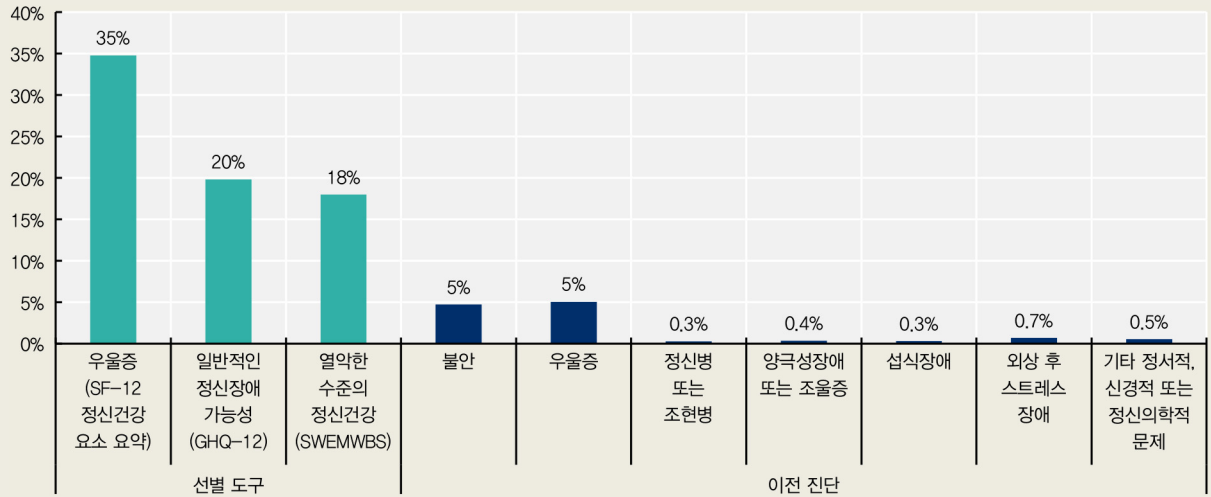
출처: 캐나다의 구조화된 면담 데이터는 캐나다 통계청(2013^[8]), *Canadian Community Health Survey: Mental Health, 2012*, The Daily, <https://www150.statcan.gc.ca/n1/daily-quotidien/130918/dq130918a-eng.htm>에서 인용, 캐나다의 PHQ-9 데이터는 Dobson, K. 외 (2020^[9]), “Trends in the prevalence of depression and anxiety disorders among Canadian working-age adults between 2000 and 2016”, *Health Reports*, Vol. 31/12, pp. 12-23, <https://doi.org/10.25318/82-003-X202001200002-ENG>에서 인용, 한국의 구조화된 면담 데이터는 KOSIS(n.d.^[10]), *Annual prevalence of mental disorders (adjusted for sex and age)*(데이터베이스), Korean Statistical Information Service, https://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=TX_117_2009_HB027&conn_path=12에서 인용, 한국의 PHQ-9 데이터는 KOSIS(n.d.^[11]), *Depressive disorder prevalence*(데이터베이스), National Health and Nutrition Survey, Korean Statistical Information Services, https://knhanes.kdca.go.kr/knhanes/sub01/sub01_05.do#none에서 인용, 미국의 구조화된 면담 데이터는 SAMHSA(2019^[12]), *Key Substance Use and Mental Health Indicators in the United States: Results from the 2018 National Survey on Drug Use and Health*(데이터베이스), Substance Abuse and Mental Health Services Administration, Rockville, MD, <https://www.samhsa.gov/data/sites/default/files/cbhsq-reports/NSDUH-NationalFindingsReport2018/NSDUHNationalFindingsReport2018.pdf>에서 인용, 미국의 PHQ-2 데이터는 미국 질병통제예방센터(2021^[13]), *Estimates of Mental Health Symptomatology, by Month of Interview: United States, 2019*(데이터베이스), U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, <https://www.cdc.gov/nchs/data/nhis/mental-health-monthly-508.pdf>에서 인용됨.

StatLink  <https://stat.link/q5bvk9>

두 번째로 그림 2.3은 선별 도구에 대한 응답을 기반으로 특정 정신장애로 진단을 받은 적이 있다고 답한 인구 비율이 정신건강 상태가 위험한 상태로 간주되는 비율보다 훨씬 낮다는 것을 보여준다. 이는 건강을 추구하는 행동에 영향을 미치는 낙인 및 정신건강에 대한 이해 부족과 함께 의료 서비스에 대한 경제성 및 접근성의 기능에 기인하는 경우가 많다. 그림의 왼쪽에는 긍정적 정신건강 수준이 낮고 심리적 고통의 위험군에 속하는 응답자 비율이 표시되며, (1) 단축형 정신건강 요소 요약(SF-12)의 채점 기준에

따른 우울증 위험군, (2) 일반 건강 설문지-12(GHQ-12)로 측정된 일반적인 정신장애 가능성이 있는 위험군, (3) 단축형 워릭-에딘버러 웰빙지수(SWEMWBS)의 채점 기준에 따라 정신적 웰빙 수준이 열악한 군도 표시된다(세 가지 도구에 대한 자세한 내용은 표 2.5, 표 2.10 및 부속서 2.B 참조). 그림의 오른쪽은 특정 정신건강 상태로 진단을 받은 적이 있다고 답한 응답자의 비율을 보여준다.

그림 2.3. 선별 도구에 의해 심리적 고통을 경험한 적이 있다고 확인된 응답자의 비율은 심리적 고통을 경험한 적이 있다고 확인된 응답자의 비율보다 훨씬 낮은 것으로 나타난다



주: 각 선별 도구에 대한 채점 정보는 다음과 같다. 우울증 위험은 SF-12 정신건강 요약 복합 척도를 변형하여 45점 이하 (0점은 최악의 정신건강, 100점은 최상의 정신건강을 나타냄)로 정의되며, 일반적인 정신장애(CMD) 가능성 위험은 GHQ-12에서 4점 이상으로 정의되고(Woodhead 외, 2012_[14]) 열악한 수준의 정신건강은 SWEMWBS 점수가 표본의 평균보다 1 표준 편차 이상 높은 것으로 정의된다. 개별 선별 도구에 대한 자세한 내용은 부속서 2.B를 참조한다.

출처: 예식스대학교 사회경제연구소(2022_[15]), *Understanding Society: Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009* (데이터베이스), 15th Edition, UK Data Service, SN: 6614, <http://doi.org/10.5255/UKDA-SN-6614-16>, from wave 10 only(2018년 1월 - 2020년 5월)를 기반으로 OECD에서 계산함.

OECD 국가에서 기존에 수집하고 있는 인구 정신건강 데이터는 무엇인가?

2022년 2월과 3월, 정신건강 결과 측정과 관련하여 OECD 국가에서 수행하고 있는 정책 및 조치를 파악하기 위해 설문조사를 실시하였다. 설문조사는 OECD 사무국이 설계한 설문지를 사용하였으며, OECD 38개 국가 중 37개국에서 답변을 제공했다³. 이 설문지에는 사용된 통계 도구(진단, 증상 경험, 선별 도구 및 구조화된 면담에 대한 문항)와 측정 결과(정신건강 문제, 긍정적 정신건강 및 외로움, 스트레스, 정신건강에 대한 태도 등 기타 관련 주제)에 대한 내용을 포함하고 있다. 의료 서비스 이용 및 의료 시스템 접근성과 관련된 정신건강 데이터에 대한 논의가 *새로운 정신건강 시스템 벤치마크*(OECD, 2021_[11])에 제시되어 있으며, 설문조사는 의료 서비스 이용이나 의료 시스템 접근성 보다는 정신건강 *결과*, 특히 행정적 데이터보다는 가구 조사를 통해 측정된 결과에 중점을 둔 기존 작업을 기반으로 하고 있다.

모든 OECD 국가는 이미 인구 정신건강에 대한 행정적 데이터와 설문조사 데이터를 수집하고 있다

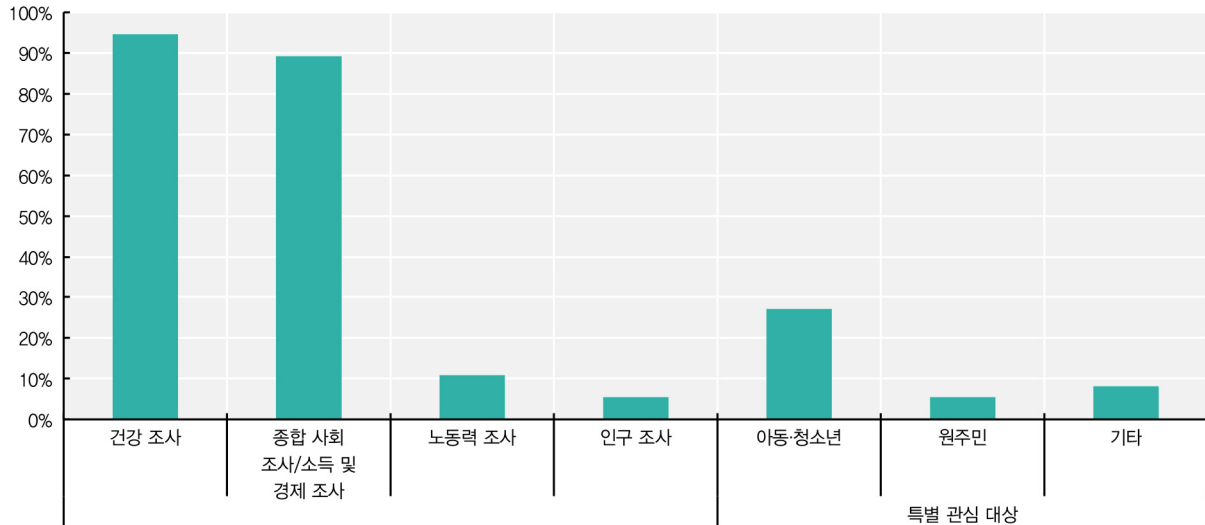
모든 OECD 국가는 자살률과 알코올 및 약물 과다복용으로 인한 사망을 포함하여 사망 원인 통계를 수집한다. 사망 원인 통계는 일반적으로 병원이나 의료 서비스 제공자가 수집하며, 경찰청은 자살로 인한 사망을 보고한다. OECD는 이미 회원국을 대상으로 자살로 인한 사망과 기타 유형의 절망사에 대한 통계를 정기적으로 발표하고 있다(OECD, 2020_[16], OECD, 2021_[17])⁴.

정신건강에 관한 행정적 데이터는 사망 기록 그 이상의 의미가 있다. 국가에 따라 병원 퇴원 등록부(Hospital discharge registries)에 의료 전문 분야별 입원 및 퇴원 기간 정보를 포함하는 방안이 캐나다, 칠레, 헝가리, 이탈리아, 슬로베니아, 스위스, 튀르키예를 포함한 여러 국가에서 논의되었다. 스페인과 영국을 포함한 일부 국가에서는 특정 행동장애의 유병률과 발병률을 측정하기 위해 치료 데이터 또는 임상 치료 데이터를 수집한다. 또한 스웨덴 사회보험청(Swedish Social Insurance Agency)에서는 정신의학적인 진단에 따른 병가를 특별 분류하여 결근의 원인에 대한 데이터를 수집한다. 마지막으로 소수의 국가에서는 정신과 약물에 대한 행정적 데이터를 수집한다. 예를 들어 프랑스에서는 *국립의약품 건강제품안전청*(Agence nationale de sécurité du médicament, ANSM)이 외래 환자에게 제공되는 항정신성 의약품에 대한 데이터를 발표하고, 네덜란드 통계청은 ATC(anatomical therapeutic chemical) 코드에 따라 정신건강 상태와 관련된 의약품을 포함하여 조제 의약품에 대한 데이터를 제공한다. 호주에서는 의료 서비스 및 의약품 보조 혜택(Pharmaceutical Benefits Scheme)에 따라 보장되는 조제 의약품에 대한 행정적 데이터를 수집한다. 슬로베니아 국립공중보건연구소(Slovenian National Institute of Public Health, NIJZ)는 정신건강 관련 의약품을 포함하여 처방 의약품 청구에 대한 데이터를 수집한다.

또한 설문조사에 응답한 모든 OECD 국가는 코로나19 팬데믹 이전부터 이미 가구 조사를 통해 정신건강 결과에 대한 인구 전체 데이터를 수집하고 있다고 보고했다. 이러한 데이터는 대부분 건강 면담 조사를 통해 수집되지만, 일반적인 사회 조사를 통해서도 정신건강 데이터를 수집한다고 보고한 국가도 89%에 달했다(그림 2.4). 정신건강에 대한 일부 데이터는 노동력 조사와 국가 인구 조사의 특별 모듈을 통해서도 수집된다. 일부 국가에서는 원주민, 형사사건에 연루된 자, 아동 및 청소년층을 포함한 하위 인구 집단을 대상으로 특별 조사를 실시하여 정신건강 데이터를 수집한다고 보고했다(후자에 대한 자세한 내용은 상자 2.2 참조).

그림 2.4. 대부분의 OECD 국가는 건강 조사 및 일반 사회 조사를 통해 인구 정신건강을 측정한다고 보고한다

인구 정신건강에 관한 설문조사에 응답한 OECD 국가의 비율



주: 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다.
출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

StatLink  <https://stat.link/40bmf6>

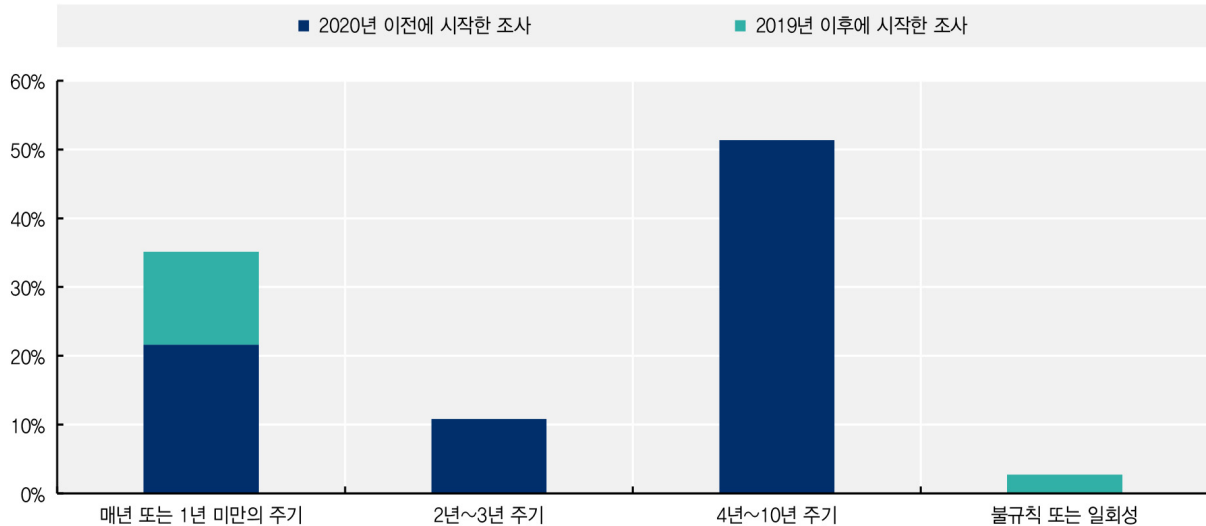
코로나19 출현 이후 많은 국가에서 정신건강 관련 설문조사를 실시했지만 향후에도 이러한 조사가 계속될지는 미지수이다

정신건강은 팬데믹으로 인해 많은 OECD 국가에서 국가적 의제로 부상했다. 그 결과 OECD 설문지에 응답한 대부분의 국가는 2020년 3월 이후 정신건강 데이터 수집 관행을 강화했다고 보고했다. OECD 국가의 약 68%는 신규 정신건강 설문조사를 단독으로 실시(43%)하거나, 기존 설문조사에 정신건강 및 코로나19 모듈을 추가하여 실시(35%)하여 팬데믹 기간에 정신건강 데이터를 추가로 수집했다고 보고했다(표 2.3 참조)⁵. 신규 정신건강 설문조사의 대부분은 매주, 격주, 매월 또는 매분기 응답자를 면담하는 높은 빈도의 설문조사에 해당한다. 그러나 이러한 설문조사가 앞으로도 동일한 빈도로 지속될지는 불분명하다. 실제로 일부 국가에서는 이미 코로나19 관련 설문조사를 중단했으며, 매주 또는 매월 실시하던 설문조사도 이후 격주 또는 분기별로 실시하는 등 그 빈도가 감소했다.

2020년 이전에는 매년 또는 1년 미만의 주기로 설문조사를 실시하여 정신건강 데이터를 수집한 국가는 22%에 불과했으며, 11%는 2년~3년마다 설문조사를 실시한 것으로 나타났다. 팬데믹 이전의 일상으로 돌아간다면 OECD 국가의 절반 이상(51%)이 4년~10년에 한 번 주기로 정신건강 데이터를 수집한다는 의미이다. 이처럼 조사 주기가 길면(코로나 19 사망자가 급증하고 엄격한 격리 조치가 시행됐던 민감한) 코로나 19 팬데믹 기간 동안 봐왔던 인구 수준의 변화를 추적하기 어렵고, 그에 따른 정책적 개입이 더욱 어려워진다.

그림 2.5. 많은 OECD 국가에서는 정신건강 데이터를 수집하는 빈도가 낮으며 절반 이상이 4년~10년에 한 번 주기로 설문조사를 실시하는 것으로 보고했다

인구 정신건강에 관한 설문조사에 응답한 OECD 국가의 비율



주: 이 그림은 각 국가의 정신건강 데이터가 포함된 전체 설문조사가 아닌 국가별로 가장 빈번하게 실시된 설문조사만을 고려한 것이다. 따라서 국가별로 가장 높은 빈도로 실시된 정신건강 설문조사의 결과를 보여준다. 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다.

출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

StatLink <https://stat.link/mboi94>

상자 2.2. 아동·청소년의 정신건강 데이터 수집을 위한 계획

코로나19 팬데믹 기간에 아동 및 청소년의 정신건강은 극심하게 악화되었으며(OECD, 2021^[18], OECD, 2021^[19]), 여러 OECD 국가에서는 자살률, 불안, 우울증 및 일반적인 심리적 스트레스 증가에 대처하기 위해 2021년과 2022년에 아동 및 청소년 정신건강을 위한 캠페인을 시작했다(HHS, 2021^[20], Chile, 2021^[21], Santé Publique France, 2021^[22]). OECD 설문조사 결과에 따르면 팬데믹으로 인해 아동 및 청소년의 정신건강에 대한 중요성이 강조되기는 했지만, 이전부터 여러 OECD 국가에서 이미 정신건강 모듈이 포함된 아동 또는 청소년 관련 설문조사를 시행하고 있었다(표 2.2).

아동 및 청소년의 정신건강 측정은 성인의 정신건강 측정과 여러 가지 면에서 차이가 있다. 일부 설문조사에서는 이전 진단에 대한 문항, WHO-5와 같은 표준화된 복합 척도, 부정적 정서에 대한 문항 등 아동 및 성인에게 동일한 도구를 사용하지만, 청소년에게 특화된 검증된 선별 도구를 사용하는 경우도 있다. OECD 설문지에 응답한 여러 국가에서는 3세~16세 아동 및 청소년을 위한 행동 선별 도구인 강점·어려움 설문지(Strengths and Difficulties Questionnaire, SDQ) 또는 발달 및 웰빙 평가(Development and Well-Being Assessment, DAWMA)를 사용하여 2세 이상 아동의 정신의학적 진단을 선별했다고 보고했다. 아동 및 청소년 설문조사에는 행동 및 정서적 문제, 아동기의 부정적 경험, 아동기의 긍정적 경험 및 약물 사용 또는 남용을 포괄하는 모듈이 포함되며 아동, 부모 또는 교사를 대상으로 한 문항이 포함될 수 있다(표 2.11). 일부 설문조사에서는 주의력결핍 과잉행동장애(ADHD) 또는 자폐 스펙트럼 장애(ASD)의 이전 진단에 대해서도 다루고 있다.

표 2.2. 여러 국가에서 아동 및 청소년의 정신건강에 대한 설문조사 또는 설문조사 모듈을 도입했다

| 국가 | 설문조사 |
|--|--|
| 호주 | 호주 아동 및 청소년 정신건강과 웰빙에 관한 조사(Australian Child and Adolescent Survey of Mental Health and Wellbeing) |
| 캐나다 | 캐나다 아동 및 청소년 건강 조사(Canadian Health Survey of Children and Youth, CHSCY) |
| 독일 | 독일 어린이 및 청소년의 건강 연구 (Study on the Health of Children and Adolescents in Germany, KiGGS) |
| 이탈리아 | 어린이 및 청소년의 삶의 질 조사(Quality of Life in Children and Adolescents)* |
| 룩셈부르크 | 룩셈부르크 청소년 조사(Youth Survey Luxembourg) |
| 영국 | 아동 및 청소년의 정신건강에 관한 조사(Mental Health of Children and Young People Surveys) |
| 미국 | 청소년 건강위험행위 조사(Youth Risk Behavior Survey, YRBS) 국민건강면접조사(National Health Interview Survey, NHIS)† |
| 덴마크, 아일랜드, 이탈리아, 라트비아, 룩셈부르크, 슬로베니아, 스웨덴 | 학령기 아동의 건강행동조사(Health Behaviour in School-Aged Children, HBSC) |

주: HBSC는 가구 조사가 아닌 학교 기반 설문조사이다. * 팬데믹이 시작된 이후(2020년 3월 이후) 도입된 설문조사임을 나타낸다. † NHIS는 핵심 모듈의 아동에 대한 설문 구성에 강점·어려움 설문지(SDQ)를 포함한다. 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다.

출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

표 2.3. OECD 국가의 절반 이상이 코로나19 팬데믹 기간에 정신건강 데이터의 수집 빈도를 증가시켰다고 보고했다

| 국가 | 단독 코로나 조사 | 기존 설문조사에 코로나 모듈 추가 | 모든 코로나 관련 조사 |
|-------|-----------|--------------------|--------------|
| 호주 | ● | | ● |
| 오스트리아 | | | |
| 벨기에 | ● | | ● |
| 캐나다 | ● | | ● |
| 칠레 | ● | | ● |
| 콜롬비아 | ● | | ● |
| 코스타리카 | | ● | ● |
| 체코 | | | |
| 덴마크 | | | |
| 핀란드 | | ● | ● |
| 프랑스 | ● | ● | ● |
| 독일 | ● | ● | ● |
| 그리스 | | | |
| 헝가리 | | | |
| 아이슬란드 | | ● | ● |
| 아일랜드 | ● | | ● |
| 이스라엘 | ● | | ● |
| 이탈리아 | | ● | ● |
| 일본 | | | |
| 한국 | ● | | ● |
| 라트비아 | | | |
| 리투아니아 | | | |
| 룩셈부르크 | ● | | ● |

| 국가 | 단독 코로나 조사 | 기존 설문조사에 코로나 모듈 추가 | 모든 코로나 관련 조사 |
|-------|-----------|--------------------|--------------|
| 멕시코 | ● | | ● |
| 네덜란드 | | ● | ● |
| 뉴질랜드 | | ● | ● |
| 노르웨이 | | ● | ● |
| 폴란드 | | | |
| 포르투갈 | | | |
| 슬로바키아 | | | |
| 슬로베니아 | ● | | ● |
| 스페인 | | ● | ● |
| 스웨덴 | ● | ● | ● |
| 스위스 | ● | | ● |
| 튀르키예 | | | |
| 영국 | | ● | ● |
| 미국 | ● | ● | ● |

주: 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다.

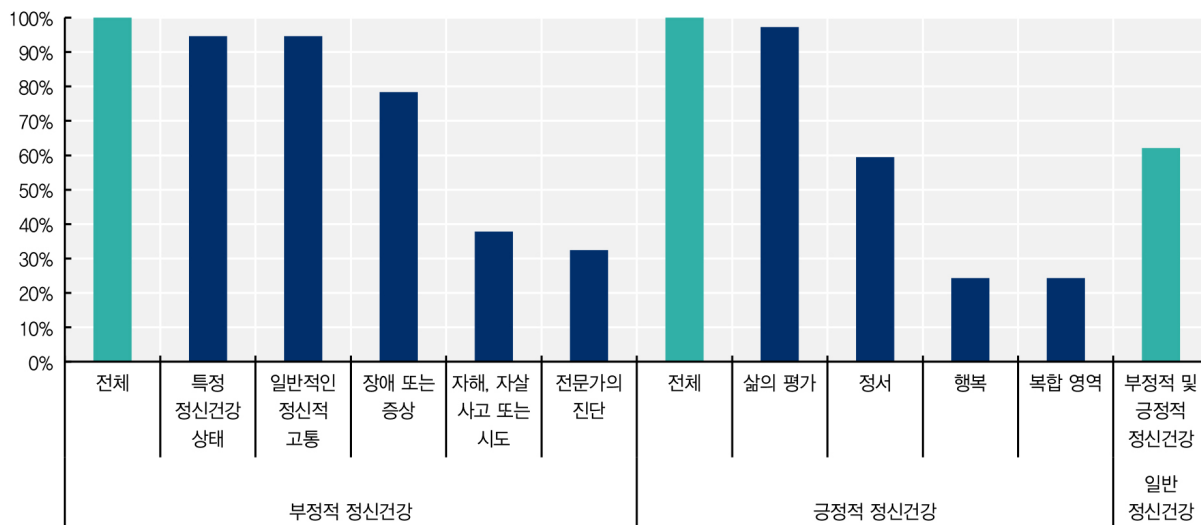
출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

가구 조사는 주로 정신건강 문제에 초점을 맞추고 있다

모든 OECD 국가는 정신건강 문제 및 긍정적 정신건강 결과에 대한 데이터를 수집한다. 정신건강 문제의 경우 사용된 도구와 측정된 결과가 매우 다양한 반면 긍정적 정신건강의 경우 국가 간 비교 데이터는 주로 생애 평가 측정에 국한되어 있다(그림 2.6). 59%의 국가에서 정서에 대한 데이터를 수집한다고 보고했고, 행복에 대한 데이터를 수집한다고 보고한 국가는 24%에 불과했다.

그림 2.6. 모든 OECD 국가에서 정신건강 문제 및 긍정적 정신건강에 대한 데이터를 수집한다고 응답했으며 후자는 주로 생애 평가에 중점을 두고 있다

인구 정신건강에 관한 설문조사에 응답한 OECD 국가 중 인구의 다양한 정신건강 결과에 대한 데이터를 수집한다고 보고한 국가의 비율



주: 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다. EU-SILC 2013조사의 특별 웰빙 모듈에서 수집된 문항은 응답자가 자신의 삶이 가치 있다고 느끼는 정도에 관한 것으로 후속 웰빙 모듈에서 삭제되었기 때문에 그림에 포함되지 않았다.

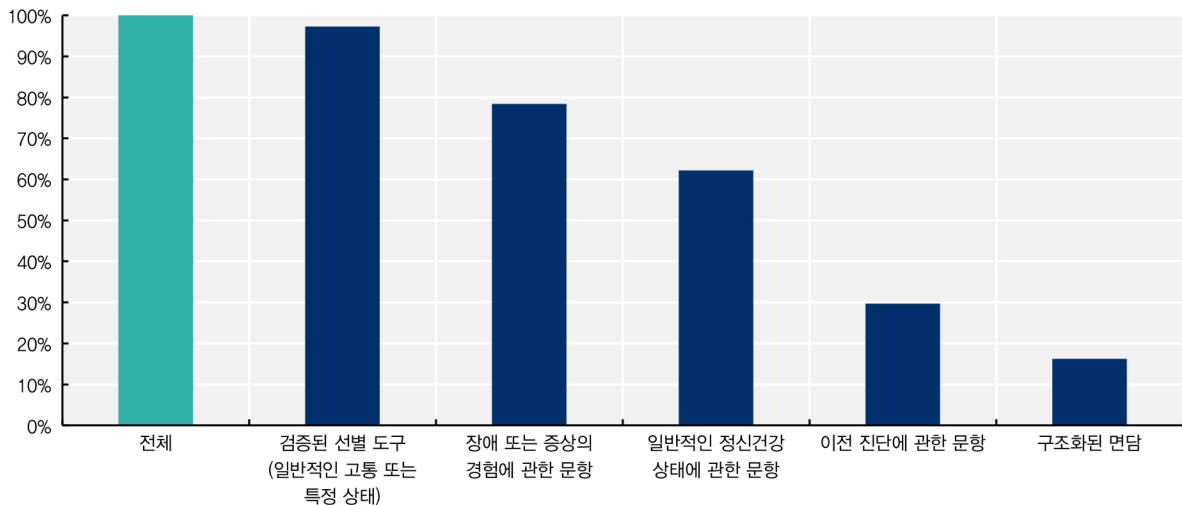
출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

StatLink <https://stat.link/cghuny>

정신건강 문제의 결과는 다양한 도구를 통해 측정된다. OECD 국가에서 가장 자주 사용되는 두 가지 도구는 선별 도구와 경험한 증상 또는 장애(일반적 또는 특정)에 대한 문항으로, 가구 조사에서 해당 유형의 도구를 사용한다고 보고한 OECD 국가는 각각 97%와 78%로 나타났다(그림 2.7). OECD 국가의 절반 이상(62%)이 사람들의 일반적인 정신건강 상태에 대해 단일 문항을 사용한다고 응답했다. 가구 조사(30%) 또는 구조화된 면담(16%)을 통해 이전 진단에 대한 데이터를 수집한다고 보고한 국가는 이보다 훨씬 적었다.


그림 2.7. OECD 국가는 선별 도구와 증상 및 장애 경험에 관한 질문이 정신건강 문제를 측정하는데 가장 일반적으로 사용하는 도구라고 보고했다

정신건강 문제를 측정하는 도구 유형별 OECD 국가 응답 비율



주: 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다.

출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

StatLink  <https://stat.link/5twjx3>

일반적인 심리적 고통과 우울증의 증상은 표준화된 선별 도구로 포착되는 경향이 있는 반면, 불안 경험에 대한 측정 은 국가 간에 통일되지 않은 경우가 많다

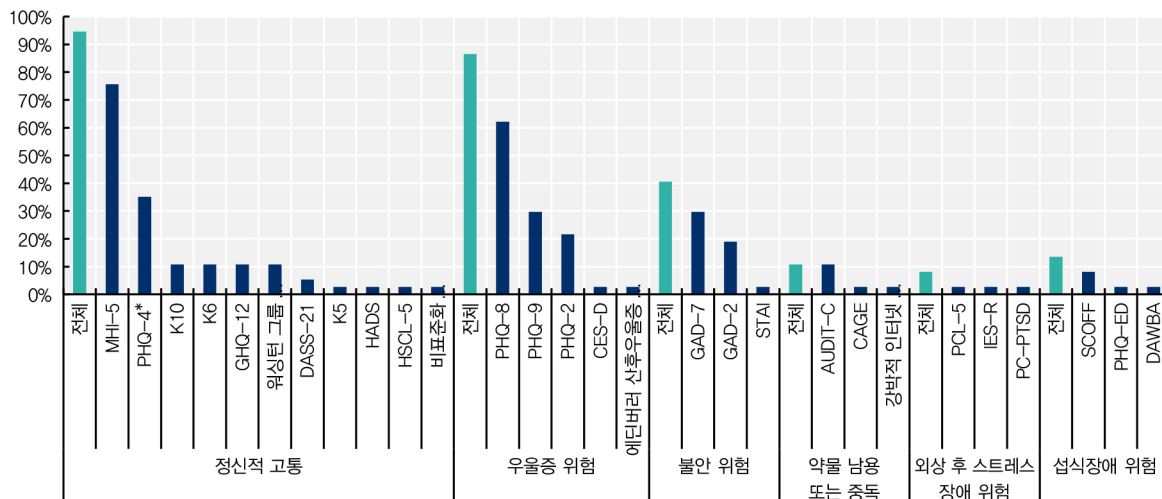
정신건강 문제의 연속체 내에서 기존의 측정 계획은 일부 유형의 정신건강 문제에 더 중점을 둔다. 불안 및 우울장애는 OECD 국가에서 사람들에게 영향을 미치는 가장 일반적인 정신건강 상태이다(OECD/European Union, 2018_[23]). 86%의 국가(37개국 중 32개국)가 우울증 증상을 측정하는 타당화된 전용 선별 도구를 보유하고 있고, 95%의 국가가 일반적인 심리적 고통(37개국 중 35개국)을 측정하는 선별 도구를 보유하고 있다. 반면에 41%의 국가(37개국 중 15개국)만이 불안 증상에 대한 선별 도구를 사용하고 있는 것으로 나타났다(그림 2.8). 국가별로 사용되는 선별 도구의 항목은 2문항에서 40문항까지 매우 다양하게 구성된다(표 2.5 참조).

우울증 증상을 측정하는 가장 일반적인 선별 도구로 PHQ를 변형하여 사용하고 있으며, OECD 국가의 84%(37개국 중 31개국)가 이를 사용하는 것으로 나타났다. MHI-5는 일반적인 심리적 고통을 측정하는 가장 일반적인 선별 도구로, OECD 국가의 76%(37개국 중 28개국)에서 사용하고 있다. 두 경우 모두 유럽연합 통계청(Eurostat)의 주도하에 사용되며, 이는 유럽연합 회원국의 데이터 수집 절차 및 방법을 통일하는 데 기여한다. MHI-5를 사용하는 28개국 중 26개국(호주와 이스라엘 제외)이 Eurostat에 속해 있다.⁶ PHQ-8은

5년~6년마다 실시되는 Eurostat의 유럽건강면접조사(European Health Interview Survey, EHIS)에 포함되었다. 여러 비유럽 OECD 국가에서도 PHQ를 변형하여 사용하고 있다(표 2.5 참조).

그림 2.8. 일반적인 심리적 고통과 우울증 증상을 포착하는 선별 도구가 불안이나 기타 장애의 증상을 포착하는 선별도구에 비해 더 일반적으로 사용된다

가구 조사에서 정신건강 문제 위험을 측정하기 위한 타당화된 선별 도구별 OECD 국가의 사용 비율



주: 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다. Eurostat의 주도하에 MHI-5 및 PHQ-8 선별 도구가 사용되며, 다른 여러 비유럽 OECD 국가에서도 해당 선별 도구, 특히 PHQ-8를 사용하는 것으로 나타났다. MHI-5는 향후 EU-SILC 특별 웰빙 모듈에서 반복적으로 사용되지 않으므로 이를 통해 데이터를 정기적으로 수집하는 국가의 비율이 감소할 것이다.

출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

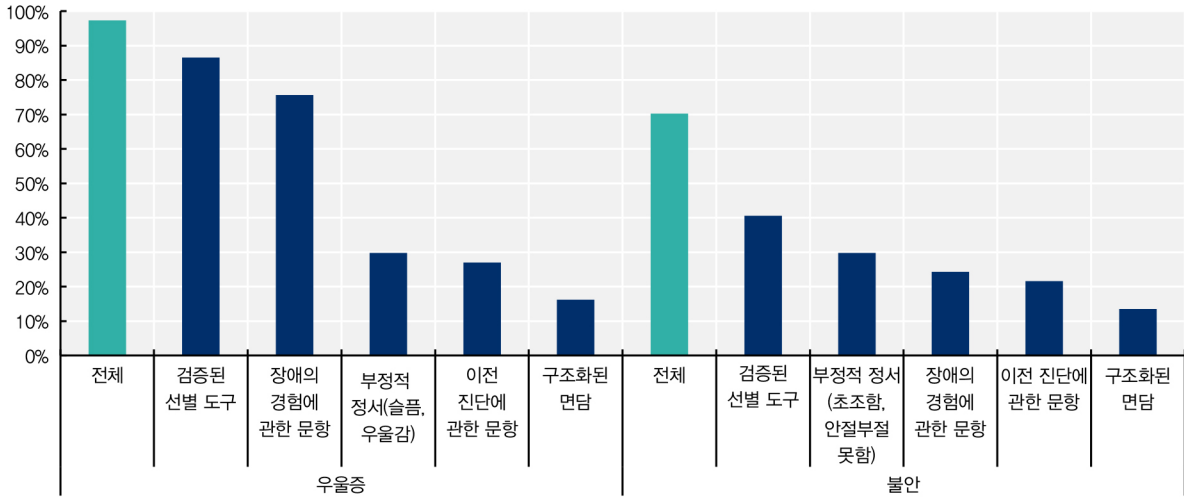
StatLink <https://stat.link/exyhv4>

또한 OECD 국가에서 불안 증상에 대한 데이터를 수집했으나 타당화된 선별 도구가 아닌 각국에 특화된 도구를 사용하여 수집된 경우가 많았다(그림 2.9). 70%의 국가에서 불안 결과를 수집하는 데 구조화된 면담, 이전 진단 또는 불안장애의 경험에 대한 문항, 정서 데이터 또는 타당화된 선별 도구를 조합하여 사용한다고 보고했다. 설문조사에 포함된 모든 측정 도구는 우울증 증상을 측정하는 데 주로 사용한다고 응답한 국가의 비율이 더 높게 나타났다. 유일한 예외는 부정적 정서에 관한 문항으로, 사용 용도가 균일하게 나뉘었다. 30%의 국가에서 불안(초조함, 불안감)과 우울증(우울감, 의기소침함)을 모두 측정하기 위해 부정적 정서를 사용한다고 응답했다.

우울증 및 불안장애에 대한 측정 계획에 중점을 둔다는 것은 우울증과 불안장애가 유병률이 가장 높은 정신건강 상태라는 사실(OECD/European Union, 2018_[23])과, 이는 전 세계 및 OECD 국가의 질병 부담에 크게 기여한다는 사실을 반영하고 있다(Santomauro 외, 2021_[24]). 외상 후 스트레스 장애(PTSD), 양극성장애, 섭식장애 등 다른 특정 정신건강 상태에 대한 데이터 수집 관행은 OECD 국가 간에 매우 불균일한 것으로 나타났다(그림 2.8).

그림 2.9. 국가에서는 불안 데이터를 수집하지만 표준화되지 않은 측정 방법을 사용하는 경우가 많다

가구 조사에서 우울증 또는 불안 증상을 측정하기 위한 도구 유형별 OECD 국가의 사용 비율



주: 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다.
출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

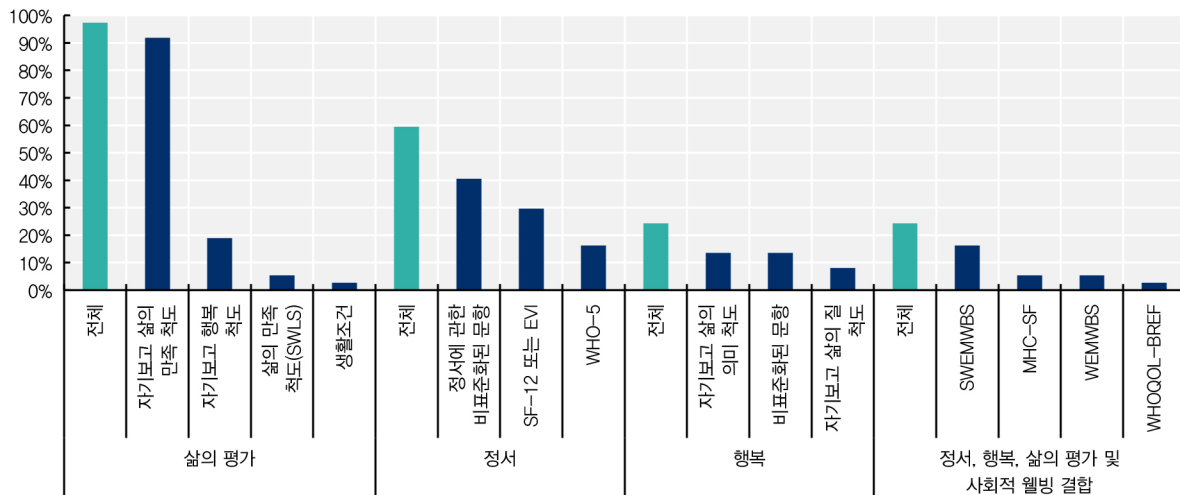
StatLink  <https://stat.link/6p1esa>

대부분의 국가에서 생애 평가 대한 비교 데이터를 수집하지만, 정서와 행복에 대해서는 수집 빈도가 낮다

거의 모든 OECD 국가에서는 생애 평가에 대한 데이터를 수집하는 데 주로 삶의 만족도에 대한 자기보고 설문조사를 통해 수집한다. 긍정적 정신건강의 다른 측면, 즉 정서 및 행복은 OECD 국가에서 실시하는 설문조사에서 다루는 빈도가 훨씬 낮고, 설령 다루는 경우에도 사용하는 도구가 국가 간에 표준화되어 있지 않다(그림 2.10). 정서 측정은 일반적으로 행복에 비해 수집되는 빈도가 높다. 59%의 국가에서 표준화된 복합 척도와 통일되지 않은 문항을 조합하여 어떤 형태로든 정서에 대한 데이터를 수집하는 반면, 행복에 대한 데이터를 수집하는 OECD 국가는 24%에 그치는 것으로 나타났다. 긍정적 정신건강 결과를 측정하기 위한 표준화된 도구로는 SF-12(및 SF-36 에너지 및 활력 지수(EVI)에 관한 하위 구성요소), WHO-5 및 WEMWBS 또는 그 단축형인 SWEMWBS 등의 세 가지 도구가 가장 일반적으로 사용되나 전반적인 사용률은 여전히 낮다. 가구 조사에서 각 척도를 사용한다고 응답한 국가는 각각 30%, 16%, 19%에 불과했다.

그림 2.10. 정서 데이터는 행복 데이터보다 수집되는 빈도가 높지만, OECD 국가에서 삶의 만족도를 넘어 긍정적 정신건강에 대한 데이터를 수집하기 위해 사용하는 도구가 통일되어 있지 않다

가구 조사에서 긍정적 웰빙을 측정하기 위한 결과 측정 도구 유형별 OECD 국가의 사용 비율



주: 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다. EU-SILC 2013 조사의 특별 웰빙 모듈에서 수집된 문항은 응답자가 자신의 삶이 가치 있다고 느끼는 정도에 관한 것으로 후속 웰빙 모듈에서 삭제되었기 때문에 그림에 포함되지 않았다.

출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

StatLink <https://stat.link/l8kq9v>

정신건강 인식에 관한 데이터 수집과 관련하여 최근 흥미로운 변화가 확인된다

전반적으로 기타 정신건강 관련 주제(예: 정신건강 약물 사용 및 서비스 이용, 가정 내 아동 및 청소년의 정신건강, 외로움 및 스트레스, 회복력 및 자기 효능감, 낙인 및 이해능력을 포함한 정신건강에 대한 태도, 충족되지 않은 요구사항에 대한 문항)에 대한 데이터 수집 관행도 국가별로 불균일하게 나타났다(표 2.11 참조). 이러한 문제 중 상당수는 해당 주제가 아직 개념적으로 명확하지 않으며, 국제적으로 표준화된 도구도 거의 없다는 점이다. 예를 들어 정신건강 낙인, 차별, 정신건강 문제나 자원에 대한 지식이나 이해 능력 등의 주제를 포괄하는 (매우 다양한) 지표를 수집한다고 보고한 국가는 30%에 불과했다⁷. 그러나 최근 정신건강 인식에 대한 관심이 높아지면서 일부 국가에서는 새로운 설문조사를 시작하고 새로운 방법을 개발했다. 예를 들어 2021년 스웨덴 공중보건연구원(Public Health Agency)은 10,000명 이상의 응답자를 대상으로 정신질환 및 자살에 대한 인식과 태도에 관한 온라인 인구 조사를 실시했다(Public Health Agency Sweden, 2022^[25]). 공중보건연구원은 기존 정신건강 낙인 측정 도구 400개 이상을 체계적으로 검토하고 인지 검사를 실시한 결과, 기존 측정의 경우 압도적으로 부정적인 어조여서 그 자체로 낙인을 씌우며 주로 중증 정신질환 사례에 초점을 맞췄다는 결론을 내렸다. 따라서 공중보건연구원은 자체 설문 조사를 개발하기로 결정했다. 최종 설문지에는 정신질환에 대한 긍정적 인식과 부정적 인식을 모두 포착하고 우울증, 불안 및 스트레스와 관련된 상태에 대한 일반적인 경험을 포함하여 모든 형태의 정신질환에 초점을 맞춘 의미분별법(단어 쌍)으로 설계된 항목이 포함되었다(Public Health Agency Sweden, 2022^[25]).

결론 및 향후 방향

인구 정신건강 결과를 측정하는 것은 OECD 국가의 공식 데이터 생산기관에 새로운 분야가 아니며, 여러 국가 통계청과 보건기관에서는 코로나19가 출현하기 이전부터 이미 관련 데이터를 수집해 왔다. 그러나 향후 개선의 여지가 있는 것도 분명하다.

첫째, 정신건강의 일부 측면은 다른 측면에 비해 더 빈번하게 측정되는 경향이 있고, 국가 간 측정 통일성을 개선할 여지가 있다. OECD가 공식 데이터 생산기관을 대상으로 실시한 설문조사의 결과에 따르면 기존의 데이터 수집 관행이 긍정적 정신건강뿐만 아니라 정신건강 문제 측면이 누락된 채 정신건강 결과의 전체적인 범위를 포착하지 못하고 있는 것으로 나타났다. 우울증 증상에 대한 선별 도구를 사용하는 국가는 86%, 일반적인 심리적 고통에 대한 선별 도구를 사용하는 국가는 95%로 나타난 반면, 불안장애, 즉 기분장애와 함께 OECD 국가에서 가장 일반적인 정신질환 중 하나인 범불안장애에 대해 표준화된 선별 도구를 사용하는 국가는 41%에 불과한 것으로 나타났다. 외상 후 스트레스 장애, 양극성장애, 섭식장애 등과 같은 다른 특정 정신건강 상태에 대한 데이터 수집 관행은 국가 간에 매우 불균일한 것으로 나타났다. 긍정적 정신건강의 경우 거의 모든 국가에서 생애 평가 데이터를 일부 수집하지만, 정서 및 행복에 대한 정보는 수집 빈도가 훨씬 낮고(각각 OECD 국가의 59% 및 24%) 표준화된 방식으로 수집되지 않는 경우가 많다. 따라서 데이터 생산기관은 첫 번째 단계로 우울증뿐만 아니라 불안 증상을 포함한 정신건강 문제의 결과 측정에 선별 도구 사용을 확장하고 긍정적 정신건강의 정서적 측면과 행복 측면을 측정하는 데 있어 통일하는 방향으로 나아갈 수 있다.

둘째, 정신건강 결과를 정기적으로 측정하고 팬데믹 초기 2년 동안 정신건강 모듈을 사용한 설문조사를 높은 빈도로 실시했던 모멘텀을 앞으로도 일부 유지해 나가는 것이 중요하다. 데이터 생성기관은 정신건강 결과를 측정하기 위한 다양한 도구를 선택할 때, 응답 부담과 정확성 사이의 상충관계가 있다는 점을 고려하여 일반적인 정신건강 상태에 대한 단일 문항을 고빈도로 실시하는 인구 조사에 추가하면 해당 정보를 정기적으로 수집하고 설문조사에서 수집한 데이터를 연결하는 데 도움이 될 수 있다. OECD 국가 중 절반 이상(62%)에서 이미 이러한 단일 항목을 설문조사에 포함하고 있으며, 문항의 문구는 매우 다양하다. 캐나다는 단일 항목의 자기보고 정신건강(SRMH) 지표의 개발을 일찍이 선도한 국가로 캐나다가 개발한 문항 구성은 이미 칠레와 독일에서 채택하여 향후 다른 국가에서도 유용한 모델이 될 수 있다. 의료 전문가로부터 받은 이전 진단에 관한 문항이 짧게 구성되어 있지만 의료 시스템을 이용한 대상에 주로 초점을 맞춰져 있어 건강 설문조사에서만 사용하는 것이 적합하다.

제3장에서는 이러한 측정 도구의 통계 품질에 대한 증거를 자세히 검토하고 국제적 통일성을 극대화하며, 응답 부담을 최소화하기 위해 각국이 채택할 수 있는 세 가지 구체적인 측정 도구에 대한 권장사항을 제공한다.

마지막으로 어떤 결과가 정책 입안자 또는 일반 대중에게 전달되는지 관계없이 특정 도구가 포괄하는 영역과 포괄하지 않는 영역(예: 이전 진단만 해당, 정서만 해당, 신체적 증상 포함 여부 및 포함하는 경우 구체적인 증상)을 포함하여 정신건강의 정확한 측면이 측정되고 있는지 투명성을 제공하는 것이 필수적이다. 이러한 정보는 결과의 맥락을 파악하고 결과 해석에 영향을 미칠 수 있는 한계점을 제공함으로써 측정의 투명성을 높이는 데 중요하다.

부속서 2.A. 국가별 정신건강 결과 측정 조사

표 2.4. 정신건강 상태를 모니터링하기 위한 구조화된 면담 개요

| 초점 | 도구 | 약어 | 항목 수 | 기준 기간 | 소요 시간 | 수집 국가 |
|------------------------------|--|------|---|-------|-------|------------------------------------|
| ICD-10 및 DSM-IV에 따른 정신 상태 진단 | 복합 국제진단면담도구 (Composite International Diagnostic Interview) | CIDI | 300개 이상의 증상에 관한 문항으로 구성되어 있으나 건너뛰기 규칙에 따라 모든 문항에 대해 응답하지 않음 | | 75분 | 호주, 캐나다, 칠레, 독일, 한국, 미국(우울 증상만 해당) |

주: 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다. 도구에 대한 자세한 내용은 부속서 2.B를 참조한다.

출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

표 2.5. 일반적인 정신 질환과 특정 정신건강 상태의 위험을 모니터링하기 위한 타당화된 선별 도구 개요

| 초점 | 측정 | 도구 | 약어 | 항목 수 | 기준 기간 | 수집 국가 |
|-------------|-----------------------------|--|--------|------------------|--------|--|
| 심리적 고통 | 부정적 및 긍정적 정서 | 정신건강검사-5 (Mental Health Inventory-5) | MHI-5 | 5개 | 지난 1개월 | 호주, 오스트리아, 벨기에, 체코, 덴마크, 핀란드, 프랑스, 독일, 그리스, 헝가리, 아이슬란드, 아일랜드, 이스라엘, 이탈리아, 라트비아, 리투아니아, 룩셈부르크, 네덜란드, 노르웨이, 폴란드, 포르투갈, 슬로바키아, 슬로베니아, 스페인, 스웨덴, 스위스, 튀르키예, 영국 |
| 심리적 고통 | 부정적 정서, 기능 손상 | 케슬러 척도 10 (Kessler Scale 10) | K10 | 10개 | 지난 4주 | 호주, 캐나다, 네덜란드, 뉴질랜드 |
| 심리적 고통 | 부정적 정서 | 케슬러 척도 6 (Kessler Scale 6) | K6 | 6개 | 지난 4주 | 호주, 일본, 스웨덴, 미국 |
| 심리적 고통 | 부정적 및 긍정적 정서, 신체적 증상, 기능 손상 | 일반 건강 설문지 (General Health Questionnaire) | GHQ-12 | 12개 | 최근 | 호주, 벨기에, 핀란드, 스페인, 영국 |
| 우울증 및 불안 증상 | 부정적 정서, 무쾌감증 기능 손상 | 환자 건강 설문지-4 (Patient Health Questionnaire-4) | PHQ-4 | 4개(우울 2개, 불안 2개) | 지난 2주 | 호주, 벨기에, 캐나다, 칠레, 핀란드, 프랑스, 독일, 아이슬란드, 한국, 슬로베니아, 스위스, 영국, 미국 |

| 초점 | 측정 | 도구 | 약어 | 항목 수 | 기준 기간 | 수집 국가 |
|-------------------------|--|--|----------------|------------------------------------|-------|--|
| 우울증 및 불안 증상 | 부정적 및 긍정적 정서, 무쾌감증 | 병원 불안-우울증 척도 (Hospital Anxiety and Depression Scale) | HADS | 14개 (우울 7개, 불안 7개) | 지난 1주 | 프랑스 |
| 우울증 및 불안 증상 | 부정적 정서 | 홉킨스 증상 체크리스트 (Hopkins Symptom Checklist) | HSCL-5 | 5 | 지난 1주 | 노르웨이 |
| 우울증 및 불안 증상 | 부정적 정서, 무쾌감증, 신체적 증상, 기능 손상 | 우울, 불안 및 스트레스 척도 (Depression, Anxiety and Stress Scale) | DASS-21 | 21개(우울 7개, 불안 7개, 비특이적 만성 스트레스 7개) | 지난 1주 | 호주, 이탈리아 |
| 일반 및 장애 인구의 우울증 및 불안 증상 | 부정적 정서, 기능 손상 | 장애 통계 위한 워싱턴 그룹의 기능에 대한 간이형 설문자-강화 (Washington Group on Disability Statistics Short Set on Functioning-Enhanced) | WG-SS Enhanced | 12개(우울 2개, 불안 2개) | 평소 | 호주, 캐나다, 뉴질랜드, 미국 |
| 일반 및 장애 인구의 우울증 및 불안 증상 | 부정적 정서, 기능 손상 | 기능에 관한 워싱턴 그룹 확장형 설문지 (Washington Group Extended Set on Functioning) | WG-ES | 37개(우울 3개, 불안 3개) | 평소 | 미국 |
| 우울 증상 | 부정적 정서, 무쾌감증, 신체적 증상, 기능 손상 (DSM-IV 및 DSM-5 기준에 따라 주요우울장애와 일치) | 환자 건강 설문지-8 (Patient Health Questionnaire-8) | PHQ-8 | 8개 | 지난 2주 | 오스트리아, 체코, 덴마크, 독일, 그리스, 헝가리, 아이슬란드, 아일랜드, 이탈리아, 라트비아, 리투아니아, 룩셈부르크, 네덜란드, 노르웨이, 폴란드, 포르투갈, 슬로바키아, 슬로베니아, 스페인, 스웨덴, 튀르키예, 영국, 미국 |
| 우울 증상 | 부정적 정서, 무쾌감증, 신체적 증상, 기능 손상 (DSM-IV 및 DSM-5 기준에 따라 주요우울장애와 일치) | 환자 건강 설문지-9 (Patient Health Questionnaire-9) | PHQ-9 | 9개 (PHQ-8 + 자살 사고에 대한 문항 1개) | 지난 2주 | 호주, 벨기에, 캐나다, 핀란드, 프랑스, 독일, 이탈리아, 한국, 슬로베니아, 스위스, 미국 |
| 우울 증상 | 부정적 정서, 무쾌감증 | 환자 건강 설문지-2 (Patient Health Questionnaire-2) | PHQ-2 | 2개 | 지난 2주 | 호주, 캐나다, 칠레, 핀란드, 독일, 이탈리아, 노르웨이, 미국 |
| 우울 증상 | 부정적 및 긍정적 정서, 무쾌감증, 신체적 증상, 기능 손상, 대인관계 문제 | 역학연구우울 척도 (Epidemiological Studies Depression Scale) | CES-D | 20개 | 지난 1주 | 멕시코 |
| 최근 산모의 우울증 증상 | 부정적 및 긍정적 정서, 무쾌감증, 기능 손상 | 에딘버러 산후우울증 척도 (Edinburg Post-natal Depression Scale) | EPDS | 6 | 지난 1주 | 이탈리아 |
| 불안 증상 | 부정적 정서, 신체적 증상, 기능 손상 | 범불안장애 척도-7 (Generalised Anxiety Disorder-7) | GAD-7 | 7개 | 지난 2주 | 호주, 벨기에, 캐나다, 핀란드, 프랑스, 독일, 아이슬란드, 한국, 슬로베니아, 스위스, 미국 |

| 초점 | 측정 | 도구 | 약어 | 항목 수 | 기준 기간 | 수집 국가 |
|------------------------|---|---|----------------|-------------------------|---------------------|------------------------------|
| 불안 증상 | 부정적 정서, 기능 손상 | 범불안장애 척도-2 (Generalised Anxiety Disorder-2) | GAD-2 | 2개 | 지난 2주 | 호주, 캐나다, 칠레, 독일, 멕시코, 영국, 미국 |
| 불안 증상 | 부정적 정서(공황과 같은 불안 포함), 기능 손상, 주관적 웰빙 | 상태-특성 불안 척도 (The State and Trait Anxiety Scale) | STAI | 40개(상태불안 20개, 특성불안 20개) | 상태불안: 현재의, 특성불안: 평소 | 이탈리아 |
| 공황장애의 증상 | 불안 발작의 유무 및 심각도, 신체적 증상 | 환자 건강 설문지-공황 발작(Patient Health Questionnaire-Panic Disorder) | PHQ-PD | 15개 | 지난 4주 | 독일, 스위스 |
| 외상 후 스트레스 장애(PTSD)의 증상 | PTSD 증상의 유무 및 심각도(DSM-5 기준과 일치) | DSM-5 기준 외상 후 스트레스 장애 체크리스트(PTSD Checklist for DSM-5) | PCL-5 | 20개 | 지난 4주 | 캐나다 |
| 외상 후 스트레스 장애(PTSD)의 증상 | PTSD 증상의 유무 및 심각도(DSM-5 기준과 일치) | DSM-5 기준 외상 후 스트레스 장애 선별 검사 (Primary Care PTSD Screen for DSM-5) | PC-PTSD-5 | 5개 | 지난 4주 | 스위스 |
| 외상 후 스트레스 장애(PTSD)의 증상 | PTSD 증상의 유무 및 심각도(DSM-IV 기준과 일치) | 사건 충격 척도 - 개정판 (Impact of Event Scale - Revised) | IES-R | 22개 | 지난 1주 | 이탈리아 |
| 광장공포증 증상 | 일상생활의 다양한 측면과 관련된 불안의 유무 및 심각도 | 불안 측정 도구 (Angstbarometer) | Angstbarometer | 12개 | 지난 1년 | 스위스 |
| 사회불안장애의 증상 | 사회불안장애의 증상 유무 및 심각도 | 축약형 사회공포증검사 (Mini-Social Phobia Inventory) | Mini-SPIN | 3개 | 지난 1주 | 핀란드, 스위스 |
| 약물 남용 또는 중독 증상 | 알코올 중독 증상의 유무 및 심각도 | CAGE 약물남용선별도구 (CAGE Substance Abuse Screening Tool) | CAGE | 4개 | 특정 회상 기간 없음 | 벨기에 |
| 약물 남용 또는 중독 증상 | 알코올 중독 증상의 유무 및 심각도 | 알코올 사용장애 식별 테스트-단축형 (Alcohol Use Disorders Identification Test-Concise) | AUDIT-C | 3개 | 특정 회상 기간 없음 | 칠레, 스웨덴 |
| 약물 남용 또는 중독 증상 | 알코올 중독 증상의 유무 및 심각도 | 알코올사용장애 식별 테스트 (Alcohol Use Disorders Identification Test) | AUDIT | 10개 | 특정 회상 기간 없음 | 프랑스, 스페인 |
| 약물 남용 또는 중독 증상 | 인터넷 중독 및 강박적, 병리적 또는 문제적 온라인 행동 문제의 유무 및 심각도(약물 중독 및 병적 도박에 대한 DSM-IV 기준과 일치) | 강박적 인터넷 사용 척도 (Compulsive Internet Use Scale) | CIUS | 14개 | 특정 회상 기간 없음 | 스위스 |
| 섭식장애의 증상 | 신경성 식욕부진 및 신경성 폭식증 유무 및 심각도 | SCOFF | SCOFF | 5개 | 지난 3개월 | 벨기에, 핀란드, 독일 |

| 초점 | 측정 | 도구 | 약어 | 항목 수 | 기준 기간 | 수집 국가 |
|--------------------------|--|---|--------|------|-------------|-------|
| 섭식장애의 증상 | 폭식장애(binge eating disorder), 신경성 폭식증(bulimia nervosa) 및 재발성 폭식장애(recurrent binge eating)의 유무 및 심각도 | 환자 건강 설문지-섭식장애 모듈(Patient Health Questionnaire-Eating Disorder Module) | PHQ-ED | 6개 | 지난 3개월 | 프랑스 |
| 7세~17세 아동 및 청소년의 섭식장애 증상 | 섭식장애 증상의 유무 및 심각도 | 발달 및 웰빙 평가의 선별문항- 섭식장애 모듈(Screening questions from the Development and Wellbeing Assessment - Eating Disorder Module) | DAWBA | 5개 | 특정 회상 기간 없음 | 영국 |

주: 이탤릭체로 표시된 국가는 해당 문항에 대한 데이터를 더 이상 수집하지 않는다고 명시적으로 언급한 국가이다. 굵은 글씨로 표시된 국가는 공식 설문조사에서 해당 도구가 보고되지 않았지만, OECD 사무국에서 해당 국가가 PHQ-8를 핵심 모듈로 포함하는 유럽건강면접조사(EHIS)에 참여한 것을 근거로 추가되었다. OECD 사무국은 PHQ-4에 대한 개별 국가의 보고와 관계없이 PHQ와 GAD(PHQ-4의 지표가 추출되는)를 모두 수집하는 국가를 위해 PHQ-4 국가 관행을 추가하였다. 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다. 영국의 데이터에는 국가 통계청에서 실시한 정신건강 관련 설문조사만 포함되며, 위임된 다른 부서에서 수집된 데이터는 포함되지 않는다. OECD 국가 중 최소 2개국에서 보고된 도구에 대한 자세한 내용은 부속서 2.B를 참조한다.

출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

표 2.6. 이전 진단에 관한 문항 개요

| 정신건강 상태 | 문항 구성 예시 | 응답 옵션 | 기준 기간 | 수집 국가 |
|-----------------------|--|-------|-----------------------|--|
| 정신건강 상태로 진단받은 경우 | 귀하는 의사 또는 간호사로부터 다음과 같은 장기적인 정신건강 상태로 진단받은 적이 있습니까? 목록: 정신건강 상태(우울증 또는 불안 포함) (호주) 귀하는 의료 전문가(정신과 전문의, 의사, 임상심리사)로부터 귀하의 정신건강 문제를 정신장애로 진단받은 적이 있습니까? (슬로베니아) | 예/아니오 | 평생 | 호주, 슬로베니아 |
| 기분장애(우울증 포함)로 진단받은 경우 | 귀하는 지금까지 의사로부터 다음과 같은 건강 문제 또는 질환으로 진단받은 적이 있습니까? 진단을 받은 경우 귀하는 치료를 받은 적이 있거나 치료를 받고 있습니까? 우울증 또는 불안 (칠레) 귀하는 지금까지 의사로부터 정신과적 또는 심리적 장애 또는 중독으로 진단받은 적이 있습니까? 우울증 또는 우울삽화 (프랑스) | 예/아니오 | 평생, 지난 12개월 또는 코로나 기간 | 호주, 오스트리아, 캐나다, 칠레, 코스타리카, 프랑스, 뉴질랜드, 슬로베니아, 스페인, 미국 |
| 불안장애로 진단받은 경우 | 귀하는 의사, 간호사 또는 기타 의료 전문가로부터 다음과 같은 상태로 진단받은 적이 있습니까? 불안 (호주) 의료 전문가로부터 다음과 같은 상태로 진단받은 적이 있습니까? 만성 불안 (코스타리카) 귀하는 지금까지 의사로부터 정신과적 또는 심리적 장애 또는 중독으로 진단받은 적이 있습니까? | 예/아니오 | 평생, 코로나 기간 | 호주, 칠레, 코스타리카, 프랑스, 뉴질랜드, 슬로베니아, 스페인, 미국 |

| 정신건강 상태 | 문항 구성 예시 | 응답 옵션 | 기준 기간 | 수집 국가 |
|------------------------------------|---|-------|------------|----------------------|
| | 불안장애(범불안, 공포증, 강박장애 등) (프랑스) | | | |
| 양극성장애 또는 조증으로 진단받은 경우 | 귀하는 의사로부터 조울증이라고도 하는 양극성장애로 진단받은 적이 있습니까? (뉴질랜드) | 예/아니오 | 평생, 코로나 기간 | 호주, 프랑스, 뉴질랜드, 슬로베니아 |
| 외상 후 스트레스 장애(PTSD)로 진단받은 경우 | 귀하는 외상 후 스트레스 장애(PTSD)로 진단받은 적이 있습니까? (캐나다) | 예/아니오 | 평생 | 호주, 캐나다 |
| 강박장애(OCD)로 진단받은 경우 | 귀하는 의료 전문가(정신과 전문의, 의사, 임상심리사)로부터 귀하의 정신건강 문제를 정신장애로 진단받은 적이 있습니까? 강박장애 (슬로베니아) | 예/아니오 | 평생 | 호주, 슬로베니아 |
| 조현병 또는 기타 정신병적 장애로 진단받은 경우 | 귀하는 지금까지 의사로부터 정신과적 또는 심리적 장애 또는 중독으로 진단받은 적이 있습니까? 조현병 (프랑스) | 예/아니오 | 평생, 코로나 기간 | 프랑스, 슬로베니아 |
| 인격장애로 진단받은 경우 | 귀하는 지금까지 의사로부터 정신과적 또는 심리적 장애 또는 중독으로 진단받은 적이 있습니까? 경계성 인격장애 (프랑스) | 예/아니오 | 평생, 코로나 기간 | 프랑스 |
| 광장공포증 또는 사회적 장애로 진단받은 경우 | 귀하는 의사, 간호사 또는 기타 의료 전문가로부터 다음의 정신건강 상태로 진단받은 적이 있습니까? 광장공포증 (호주) | 예/아니오 | 평생 | 호주 |
| 중독장애 또는 약물 남용 문제로 진단받은 경우 | 귀하는 의사, 간호사 또는 기타 의료 전문가로부터 다음과 같은 상태로 진단받은 적이 있습니까? 알코올 또는 약물의 유해한 사용 또는 의존증 (호주) 귀하는 지금까지 의사로부터 정신과적 또는 심리적 장애 또는 중독으로 진단받은 적이 있습니까? 중독 또는 중독장애 (프랑스) | 예/아니오 | 평생, 코로나 기간 | 호주, 프랑스 |
| 섭식장애로 진단받은 경우 | 귀하는 의료 전문가(정신과 전문의, 의사, 임상심리사)로부터 귀하의 정신건강 문제를 정신장애로 진단받은 적이 있습니까? 섭식장애 (슬로베니아) | 예/아니오 | 평생, 코로나 기간 | 프랑스, 슬로베니아 |
| 품행장애 또는 행동/정서 문제로 진단받은 경우 | 귀하는 의사, 간호사 또는 기타 의료 전문가로부터 다음과 같은 상태로 진단받은 적이 있습니까? 행동 또는 정서 문제 (호주) 귀하는 의료 전문가로부터 품행장애로 진단받은 적이 있습니까? (스페인) | 예/아니오 | 평생 | 호주, 스페인 |
| 신경다양성: 주의력결핍 과잉행동장애(ADHD)로 진단받은 경우 | 귀하는 의사 또는 기타 의료 전문가로부터 주의력결핍 과잉행동장애(ADHD) 또는 주의력결핍장애(ADD)로 진단받은 적이 있습니까? (미국) | 예/아니오 | 평생 | 독일, 미국 |
| 신경다양성: 자폐스펙트럼장애(ASD)로 진단받은 경우 | 귀하는 의료 전문가로부터 자폐증으로 진단받은 적이 있습니까? (스페인) | 예/아니오 | 평생 | 스페인 |

| 정신건강 상태 | 문항 구성 예시 | 응답 옵션 | 기준 기간 | 수집 국가 |
|---------------------|--|-------|-----------------------|--------------------------------|
| 기타 정신건강 상태로 진단받은 경우 | 귀하는 의료 전문가로부터 진단받은 다른 장기적인 신체적 또는 정신적 건강 상태가 있습니까? (캐나다) 귀하는 의료 전문가(정신과 전문의, 의사, 임상심리사)로부터 귀하의 정신건강 문제를 정신장애로 진단받은 적이 있습니까? (슬로베니아) | 예/아니오 | 평생, 지난 12개월 또는 코로나 기간 | 호주, 캐나다, 칠레, 코스타리카, 프랑스, 슬로베니아 |

주: 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다. 영국의 데이터에는 국가 통계청에서 실시한 정신건강 관련 설문조사만 포함되며, 위임된 다른 부서에서 수집된 데이터는 포함되지 않는다.
출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

표 2.7. 증상 및 정신건강 상태의 경험에 관한 문항 개요

| 정신건강 상태 | 문항 구성 예시 | 응답 옵션 | 기준 기간 | 수집 국가 |
|-----------------------------|--|-------|---------------------|--|
| 자기보고 정신건강 문제 | 귀하는 지난 3개월 동안 심리적 스트레스나 급성 질환을 경험한 적이 있습니까? (이스라엘) 귀하는 현재 정신건강 문제를 겪고 있습니까? (슬로베니아) 귀하는 정신질환을 직접 경험한 적이 있습니까? (스웨덴) 귀하는 정신건강 문제를 경험했던 적이 있습니까? (미국) | 예/아니오 | 평생, 지난 12개월, 지난 3개월 | 헝가리, 이스라엘, 슬로베니아, 스웨덴, 미국 |
| 자기보고 기분장애(우울증 등) 또는 기분장애 증상 | 귀하는 지난 12개월 동안 다음의 질병이나 상태를 경험한 적이 있습니까? 우울증 [OECD 국가 중 유럽건강면접조사(EHIS)에 참여하는 유럽 국가] 귀하는 기분장애를 겪고 있습니까? (캐나다) 다음은 현재 다양한 만성 질환 또는 건강 상태의 유무와 관한 문항입니다. 만성 질환은 유병기간이 길고 일반적으로 천천히 진행되는 질환입니다. 귀하는 만성 우울증을 겪고 있습니까? (코스타리카) | 예/아니오 | 평생, 지난 12개월, 현재 | 호주, 오스트리아, 벨기에, 캐나다, 코스타리카, 체코, 덴마크, 프랑스, 독일, 그리스, 헝가리, 아이슬란드, 아일랜드, 이탈리아, 라트비아, 리투아니아, 룩셈부르크, 네덜란드, 노르웨이, 폴란드, 포르투갈, 슬로바키아, 슬로베니아, 스페인, 스웨덴, 튀르키예, 영국, 미국 |

| 정신건강 상태 | 문항 구성 예시 | 응답 옵션 | 기준 기간 | 수집 국가 |
|-------------------------|--|-------|-------------------------|---|
| 자기보고 불안장애 또는 불안 증상 | <p>귀하는 불안장애를 겪고 있습니까?(캐나다)</p> <p>귀하는 지난 12개월 동안 다음에 나열된 만성 질환/질환을 앓았던 적이 있거나 앓고 있습니까? 불안장애(예: 공황 발작, 불안) (그리스)</p> <p>귀하는 만성 불안을 겪고 있습니까? (스페인)</p> | 예/아니오 | 평생, 지난 12개월, 지난 3개월, 현재 | 호주, 캐나다, 코스타리카, 그리스, 헝가리, 노르웨이, 슬로베니아, 스페인, 스웨덴 |
| 자기보고 양극성장애 또는 조증 | <p>귀하는 다음의 상태를 겪고 있습니까? 양극성장애 (호주)</p> <p>귀하는 우울증, 양극성장애, 조증 또는 기분부전증과 같은 기분장애를 겪고 있습니까? (캐나다)</p> | 예/아니오 | 평생 | 호주, 캐나다 |
| 자기보고 외상 후 스트레스 장애(PTSD) | <p>귀하는 현재 외상 후 스트레스 장애(PTSD)를 겪고 있습니까? (캐나다)</p> | 예/아니오 | 평생 | 호주, 캐나다 |
| 자기보고 강박장애(OCD) | <p>귀하는 다음의 상태를 겪고 있습니까? 강박장애(OCD) (호주)</p> | 예/아니오 | 평생 | 호주 |
| 자기보고 조현병 또는 기타 정신병적 장애 | <p>(귀하가 언급한 질환 외에) 다음과 같은 기타 정신건강, 행동 또는 인지적 상태를 겪고 있습니까? 조현병 (호주)</p> <p>조현병, 분열형 및 망상장애(를 겪고 있습니까?) (헝가리)</p> | 예/아니오 | 평생 또는 지난 12개월 | 호주, 헝가리 |
| 자기보고 광장공포증 또는 사회적 장애 | <p>귀하는 다음의 상태를 겪고 있습니까? 광장공포증 (호주)</p> | 예/아니오 | 평생 | 호주 |
| 자기보고 중독장애 또는 약물 남용 문제 | <p>(귀하가 언급한 질환 외에) 다음과 같은 기타 정신건강, 행동 또는 인지적 상태를 겪고 있습니까?</p> | 예/아니오 | 평생 또는 지난 12개월 | 호주, 헝가리 |

| 정신건강 상태 | 문항 구성 예시 | 응답 옵션 | 기준 기간 | 수집 국가 |
|-------------------------|---|--------------------|---------------|---------|
| | 알코올 의존증, 약물 의존증, 의약품, 처방 의약품의 유해한 사용 또는 의존증 (호주) | | | |
| 자기보고 섭식장애 | 귀하는 지난 12개월 동안 다음과 같은 행동을 얼마나 자주 했습니까? a. 더 날씬해지고 싶은 욕구에 강박적으로 몰두함 b. 체중 감량을 위해 구토를 한 적이 있음 c. 체중 관리를 위해 식습관을 바꿈 (캐나다) | 전혀 없음/몇 번/매월/매주/매일 | 지난 12개월 | 캐나다 |
| 자기보고 품행장애 또는 행동/정서 문제 | 귀하는 지난 12개월 동안 품행장애를 겪은 적이 있습니까? (스페인) | 예/아니오 | 평생 또는 지난 12개월 | 호주, 스페인 |
| 자기보고 주의력결핍 과잉행동장애(ADHD) | (귀하가 언급한 질환 외에) 다음과 같은 기타 정신건강, 행동 또는 인지적 상태를 겪고 있습니까? 주의력결핍 과잉행동장애(Attention deficit hyperactivity disorder, ADHD) (호주) | 예/아니오 | 평생 | 호주 |
| 자기보고 자폐스펙트럼 장애(ASD) | 귀하는 지난 12개월 동안 자폐증을 겪은 적이 있습니까? (스페인) | 예/아니오 | 평생 또는 지난 12개월 | 호주, 스페인 |
| 자기보고 치매 | (귀하가 언급한 질환 외에) 다음과 같은 기타 정신건강, 행동 또는 인지적 상태를 겪고 있습니까? 치매(알츠하이머병 포함) (호주) | 예/아니오 | 평생 | 호주 |
| 자기보고 지적장애 | (귀하가 언급한 질환 외에) 다음과 같은 기타 정신건강, 행동 또는 인지적 상태를 겪고 있습니까? 지적장애, 정신 지체 (호주) | 예/아니오 | 평생 | 호주 |
| 자기보고 학습장애 | (귀하가 언급한 질환 외에) 다음과 같은 기타 | 예/아니오 | 평생 | 호주 |

| 정신건강 상태 | 문항 구성 예시 | 응답 옵션 | 기준 기간 | 수집 국가 |
|--------------|---|-------|---------------|----------------|
| | 정신건강, 행동 또는 인지적 상태를 겪고 있습니까? 학습장애(난독증 포함) (호주) | | | |
| 자기보고 기타 정신장애 | (귀하가 언급한 질환 외에) 다음과 같은 기타 정신건강, 행동 또는 인지적 상태를 겪고 있습니까? 기타 정신 또는 행동 문제 (호주) | 예/아니오 | 평생 또는 지난 12개월 | 호주, 코스타리카, 헝가리 |

주: 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다. 영국의 데이터에는 국가 통계청에서 실시한 정신건강 관련 설문조사만 포함되며, 위임된 다른 부서에서 수집된 데이터는 포함되지 않는다.
출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

표 2.8. 자살 사고 및 자살 시도에 관한 문항 개요

| 정신건강 상태 | 문항 구성 예시 | 응답 옵션 | 기준 기간 | 수집 국가 |
|---------|---|----------------------|---------------------|--|
| 자살 총동 | PHQ-9의 마지막 문항 | 전혀 없음/며칠/절반 이상/거의 매일 | 지난 2주 | 호주, 벨기에, 캐나다, 핀란드, 프랑스, 독일, 이탈리아, 한국, 슬로베니아, 스위스, 미국 |
| 자살 총동 | 귀하는 코로나19 팬데믹이 시작된 이후 심각하게 자살을 고려한 적이 있습니까? (캐나다) 귀하는 지난 12개월 동안 [자살을 심각하게 고려한] 적이 있습니까? (칠레) 귀하는 지난 12개월 동안 자살하는 것에 대해 생각한 적이 있습니까? (프랑스) 귀하는 살면서 자살을 심각하게 고려했던 상황에 처한 적이 있습니까? (스웨덴) | 예/아니오 | 평생, 지난 12개월, 코로나 기간 | 호주, 벨기에, 캐나다, 칠레, 핀란드, 프랑스, 한국, 멕시코, 슬로베니아, 스웨덴, 스위스, 미국 |
| 자해 행동 | 때때로 사람들은 자살할 의도 없이 자기 신체 일부를 고의로 상하게 하는 행동을 합니다. 귀하는 지난 12개월 동안 자살할 의도 없이 자해 행동을 한 적이 있습니까? (캐나다) | 예/아니오 | 평생, 지난 12개월, 지난 2주 | 호주, 캐나다, 핀란드, 그리스, 멕시코 |

| 정신건강 상태 | 문항 구성 예시 | 응답 옵션 | 기준 기간 | 수집 국가 |
|------------------------|---|----------------------|---------------------|--|
| | 귀하는 지난 2주 동안 자해하고 싶은 생각이 얼마나 자주 들었습니까? (그리스) | 전혀 없음/며칠/절반 이상/거의 매일 | | |
| 자살 시도 | 귀하는 지난 12개월 동안 자살을 시도한 적이 있습니까? (벨기에) 귀하는 지난 12개월 동안 자살을 시도한 적이 있습니까? (한국) 귀하는 자살을 시도한 적이 있습니까? (룩셈부르크) | 예/아니오 | 평생, 지난 12개월, 코로나 기간 | 호주, 벨기에, 캐나다, 칠레, 핀란드, 프랑스, 한국, 룩셈부르크, 스웨덴, 미국 |
| 자살 시도로 인해 입원 또는 치료 여부 | 귀하는 자살을 시도하여 하루밤 이상 병원에 입원한 적이 있습니까? (미국) | 예/아니오 | 평생 또는 지난 12개월 | 호주, 미국 |
| 자살 충동 또는 자살 시도 후 상담 여부 | 귀하는 자살 충동을 느낀 후 누군가와 이야기를 나누었습니까? (스위스) 귀하는 지난 12개월 동안 자살 시도로 인해 의사 또는 기타 의료 전문가로부터 치료를 받은 적이 있습니까? (미국) | 예/아니오 | 평생 또는 지난 12개월 | 프랑스, 스위스, 미국 |

주: 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다. 영국의 데이터에는 국가 통계청에서 실시한 정신건강 관련 설문조사만 포함되며, 위임된 다른 부서에서 수집된 데이터는 포함되지 않는다.
출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

표 2.9. 일반적인 정신건강 상태에 관한 문항 개요

| 정신건강 상태 | 문항 구성 예시 | 응답 옵션 | 기준 기간 | 수집 국가 |
|-------------------|---|--|-------------|--|
| 자기보고 일반적인 정신건강 상태 | 일반적으로 귀하의 정신건강은 어떻습니까? (캐나다, 칠레, 독일) 평소 귀하의 정신 상태는 어떻습니까? (이스라엘) | 훌륭하다/매우 좋다/좋다/보통이다/나쁘다 (호주, 캐나다, 칠레, 독일) 매우 좋다/좋다/보통이다/나쁘다 (이스라엘) | 현재 또는 지난 4주 | 호주, 캐나다, 칠레, 코스타리카, 핀란드, 독일, 아이슬란드, 이스라엘, 슬로베니아, 스위스, 미국 |
| 자기보고 정신건강 일수 | 귀하는 지난 30일 동안 정신건강 상태가 좋지 않았던 적이 얼마나 자주 있었습니까? (미국) | [일수] | 지난 30일 | 미국 |
| 자기보고 회복 | 귀하는 현재 자신의 정신건강 문제가 회복 중이거나 회복되었다고 생각하십니까? (미국) | 예/아니오 | 일반 평가 | 미국 |

| 정신건강 상태 | 문항 구성 예시 | 응답 옵션 | 기준 기간 | 수집 국가 |
|----------------------------|---|---|-----------------------------|---|
| 자기보고 정신건강 상태에 대한 만족도 | <p>귀하의 정신건강에 어느 정도 만족하는지 1점부터 10점까지의 척도로 표시할 수 있습니까? (네덜란드)</p> <p>귀하의 정신건강에 얼마나 만족합니까? (노르웨이)</p> | 0점(전혀 불만족) ~ 10점(매우 만족) | 일반 평가 | 네덜란드, 노르웨이 |
| 자기보고 정신건강 상태 및 코로나19 | <p>팬데믹 이전과 비교했을 때 현재 귀하의 정신건강은 어떻습니까? (캐나다)</p> <p>“지난 12개월(또는 2020년) 동안 코로나19 팬데믹으로 인해 귀하의 정신건강/웰빙에 영향을 받은 적이 있습니까?” (덴마크, 라트비아, 슬로바키아, 슬로베니아, 튀르키예)</p> <p>팬데믹으로 인해 귀하의 정신건강에 어떤 영향을 받았습니까? (스위스)</p> <p>봉쇄 기간에 귀하의 기운을 북돋을 수 없을 정도로 심각한 좌절감을 느낀 적이 있습니까? (프랑스)</p> | <p>훨씬 나아짐/약간 나아짐/거의 비슷함/약간 나빠짐/훨씬 나빠짐 (캐나다)</p> <p>1. 예, 부정적인 영향을 받았습니다 2. 예, 긍정적인 영향을 받았습니다 3. 아니요, 영향을 받지 않았습니다 (덴마크, 라트비아, 포르투갈, 슬로바키아, 튀르키예)</p> <p>0점(매우 나빠짐) ~ 10점(매우 좋아짐) (스위스)</p> <p>예/아니요 (프랑스)</p> | 코로나19 기간 | 캐나다, 덴마크, 핀란드, 프랑스, 독일, 이스라엘, 일본, 라트비아, 네덜란드, 포르투갈, 슬로바키아, 슬로베니아, 스위스, 튀르키예 |
| 정신건강이 일상 활동에 지장을 초래(손상 일수) | <p>귀하의 정신 상태가 직장 생활에 지장을 초래합니까? 아울러 가정 생활에도 지장을 초래합니까? (이스라엘)</p> <p>귀하는 지난 12개월 동안 일상생활에 지장을 받을 정도로 심각한 슬픔이나 절망감을 2주 이상 느낀 적이 있습니까? (한국)</p> <p>귀하는 지난 12개월 동안 거의 매일 지속된 심각한 슬픔이나 절망감을 2주 이상 느껴 일상적인 활동을 중단한 적이 있습니까? (미국)</p> | 예/아니요 | 지난 12개월부터 지난 4주까지 다양함 | 호주, 캐나다, 헝가리, 이스라엘, 한국, 스페인, 미국 |

주: 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다. 영국의 데이터에는 국가 통계청에서 실시한 정신건강 관련 설문조사만 포함되며, 위임된 다른 부서에서 수집된 데이터는 포함되지 않는다.
출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

표 2.10. 긍정적 정신건강 지표 개요

| 정신건강 요소 | 도구 | 약어 | 항목 수 | 기준 기간 | 수집 국가 |
|---|--|--|-------|------------------------|---|
| 긍정적 정서 | WHO-5 웰빙지수(WHO-5 Wellbeing Index) | WHO-5 | 5개 | 지난 2주 | 프랑스, 헝가리, 이탈리아, 라트비아, 뉴질랜드, 슬로베니아 |
| 긍정적 및 부정적 정서, 기능 손상 (정신건강 요소 요약) | 단축형 건강 상태(Short Form Health Status) | SF-12 | 12개 | 지난 4주 | 칠레, 이탈리아, 네덜란드, 뉴질랜드, 스페인, 미국 |
| 긍정적 및 부정적 정서, 기능 손상 | SF-36 | SF-36 | 36개 | 지난 4주 | 호주, 독일 |
| 긍정적 및 부정적 정서 | SF-36 활력 하위 척도 | EVI | 4개 | 지난 4주 | 호주, 벨기에, 이탈리아, 스위스 |
| 긍정적 또는 부정적 정서 | 정서에 관한 비표준화된 문항 문항 예시: 귀하는 어제 하루 동안 행복감을 느꼈습니까? (프랑스) 이 기간[지난 12개월] 동안 다음과 같은 감정을 어느 정도 경험했습니까? 스트레스 및 불안 (이스라엘) 이제 일련의 감정이나 감정들을 언급하겠습니다. 귀하는 지난 2주 동안 다음과 같은 감정을 얼마나 자주 느꼈습니까? 화가남 낙관적 걱정 행복 슬픔 평온 피곤함 쓸모있음 (칠레) | NA 0점(행복함 최저치) - 10점(행복함 최고치) (프랑스) 매우 그렇다/그렇다/보 통이다/전혀 그렇지 않다(이스라엘) 전혀 없음/가끔 있음/거의 있음/항상 있음 (칠레) | 1개~8개 | 어제부터 지난 1년까지 다양함 | 칠레, 코스타리카, 핀란드, 프랑스, 아일랜드, 이스라엘, 이탈리아, 일본, 라트비아, 네덜란드, 뉴질랜드, 노르웨이, 슬로베니아, 스웨덴, 영국 |
| 행복 | 삶이 가치있거나 의미있다고 느끼는 정도에 대한 자기보고 문항 예시: 귀하는 자신의 삶에서 하는 일이 의미와 가치가 있다고 느끼니까? 0점(의미 없음)~10점(매우 의미 있음)까지의 척도로 응답 (프랑스) | NA | 1개 | 일반 평가 | 호주, 오스트리아, 벨기에, 캐나다, 체코, 덴마크, 핀란드, 프랑스, 독일, 그리스, 헝가리, 아이슬란드, 아일랜드, 이탈리아, 일본, 라트비아, 리투아니아, 룩셈부르크, 네덜란드, 뉴질랜드, 노르웨이, 폴란드, 포르투갈, 슬로바키아 공화국, 슬로베니아, 스페인, 스웨덴, 스위스, 튀르키예, 영국 |

| 정신건강 요소 | 도구 | 약어 | 항목 수 | 기준 기간 | 수집 국가 |
|---------|---|--------------------------------------|------|-------|----------------------|
| | <p>귀하는 평소 자신에 대해 어떻게 설명할 수 있습니까? 다음과 같이 응답해주세요.</p> <p>1: 행복하고 삶에 관심이 많다 2: 어느 정도 행복하다 3: 다소 불행하다 4: 불행하고 삶에 관심이 거의 없다 5: 너무 불행하고 삶이 가치가 없다 (캐나다)</p> | | | | |
| 행복 | <p>자기보고 삶의 질</p> <p>문항 예시: 귀하의 삶의 질을 다음과 같이 평가해주세요. .</p> <p>1: 훌륭하다 2: 매우 좋다 3: 좋다 4: 보통이다 5: 나쁘다 (캐나다)</p> | NA | 1개 | 일반 평가 | 캐나다, 코스타리카, 핀란드, 스위스 |
| 행복 | <p>자기보고 자기 만족도</p> <p>문항 예시: 귀하는 자신에 대해 얼마나 만족합니까? (코스타리카)</p> | 1 매우 만족, 2 만족, 3 보통, 4 불만족, 5 매우 불만족 | 1 | 일반 평가 | 코스타리카, 핀란드 |
| 행복 | <p>자기보고 목적의식, 성취감 또는 목표 달성도</p> <p>문항 예시: 지금까지 나는 인생에서 중요한 목표를 달성했다 (멕시코)</p> <p>내 인생에는 분명한 목적의식이 있다 (미국)</p> <p>나는 대부분의 날에 내가 하는 일에서 성취감을 느낀다 (미국)</p> | NA | 1개 | 일반 평가 | 멕시코, 미국 |
| 행복 | <p>자기보고 사회에 도움이 되는 구성원이라는 의식</p> | NA | 1 | 일반 평가 | 헝가리 |

| 정신건강 요소 | 도구 | 약어 | 항목 수 | 기준 기간 | 수집 국가 |
|----------------------------|---|--------|------|--------|---|
| | <p>문항 예시: 귀하는 자신이 사회의 중요하고 도움이 되는 구성원이라는 것에 대해 어떻게 생각합니까? (헝가리)</p> | | | | |
| 삶의 평가 | <p>자기보고 삶의 만족도</p> <p>문항 예시: 귀하는 전반적으로 요즘 삶에 얼마나 만족합니까? 0점~10점까지의 척도로 답해주세요. 0점은 '전혀 만족하지 않음', 10점은 '완전히 만족함'을 의미합니다. [OECD 국가 중 유럽건강면접조사(EU-SILC) 웰빙 모듈에 참여하는 유럽 국가]</p> | NA | 1 | 일반 평가 | 호주, 오스트리아 , 벨기에, 캐나다, 칠레, 체코 , 덴마크, 핀란드, 프랑스, 독일 , 그리스, 헝가리, 아이슬란드 , 아일랜드, 이스라엘, 이탈리아, 한국 , 라트비아, 리투아니아, 룩셈부르크 , 멕시코 , 네덜란드, 뉴질랜드, 노르웨이, 폴란드, 포르투갈, 슬로바키아, 슬로베니아, 스페인 , 스웨덴, 스위스, 튀르키예, 영국, 미국 |
| 삶의 평가 | <p>삶의 만족 척도(Satisfaction with Life Scale, SWLS)</p> | SWLS | 5개 | 일반 평가 | 노르웨이, 슬로베니아 |
| 삶의 평가 | <p>자기보고 행복</p> <p>문항 예시: 귀하는 전반적으로 얼마나 행복하다고 생각합니까? 1점~10점 중 하나의 상자에 체크해주세요. 1점은 매우 불행함, 10점은 매우 행복함을 의미합니다. (ISL)</p> | NA | 1개 | 일반 평가 | 칠레, 프랑스, 아이슬란드, 일본 , 네덜란드, 스위스, 미국 |
| 삶의 평가 | <p>자기보고 생활조건</p> <p>문항 예시: 현재 귀하의 가정 내 생활조건은 다음과 같습니다. 1. 아주 좋음, 2. 좋음, 3. 보통, 4. 나쁨 (콜롬비아)</p> | NA | 1개 | 일반 평가 | 콜롬비아 |
| 긍정적 정서, 행복, 삶의 만족도, 사회적 웰빙 | <p>단축형 정신건강 연속성 척도(Mental Health Continuum Short Form)</p> | MHC-SF | 14개 | 지난 1개월 | 캐나다, 슬로베니아 |
| 긍정적 정서, 행복, 사회적 웰빙 | <p>워릭-에딘버러 웰빙지수(Warwick-Edinburgh Mental Well-being Scale)</p> | WEMWBS | 14개 | 지난 2주 | 핀란드, 노르웨이* |

| 정신건강 요소 | 도구 | 약어 | 항목 수 | 기준 기간 | 수집 국가 |
|----------------------------|---|-------------|------|-------|------------------------------|
| 긍정적 정서, 행복, 사회적 웰빙 | 단축형 워릭-에딘버러 웰빙지수(Short Warwick-Edinburgh Mental Well-being Scale) | SWEMWBS | 7개 | 지난 2주 | 캐나다, 핀란드, 독일, 아이슬란드, 스웨덴, 영국 |
| 긍정적 및 부정적 정서, 행복, 자존감, 집중력 | WHO 삶의 질 간편형 척도 정신건강 영역(WHO Quality of Life-BREF psychological health domain) | WHOQOL-BREF | 6개 | 지난 2주 | 칠레 |

주: *노르웨이는 현재 WEMWBS를 사용하지 않지만, 향후 국가 삶의 질 조사에 해당 도구가 포함될 수 있다는 입장을 밝혔다. 이탤릭체로 표시된 국가는 해당 문항에 대한 데이터를 더 이상 수집하지 않는다고 명시적으로 언급한 국가이다. 굵은 글씨로 표시된 국가는 공식 설문조사에서 해당 도구가 보고되지 않았지만, OECD 사무국에서 해당 국가가 “전반적으로 귀하가 하는 일이 삶에 어느 정도 가치가 있다고 생각합니까?”라는 질문이 포함된 유럽연합 소득 및 생활조건 통계(EHIS)의 2013 특별 웰빙 모듈에 참여한 것을 근거로 추가되었으며, 2018년에 해당 척도는 다시 제외되었다. 굵은 글씨 및 이탤릭체로 표시된 국가는 공식 설문조사에서 해당 도구가 보고되지 않았지만, OECD 사무국에서 해당 국가가 2016년 주관적 웰빙 측정에 대한 OECD 설문조사에 참여한 것을 근거로 추가되었다. 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다. 영국의 데이터에는 국가 통계청에서 실시한 정신건강 관련 설문조사만 포함되며, 위임된 다른 부서에서 수집된 데이터는 포함되지 않는다. OECD 국가 중 최소 2개국에서 보고된 도구에 대한 자세한 내용은 부속서 2.B를 참조한다.

출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

표 2.11. 국가별로 수집된 기타 정신건강 관련 주제에 대한 개괄적인 개요

| 주제 영역 | 도구 유형 | 지표 유형 | 수집 국가 |
|-----------------|---------------------------|--|---|
| 정신건강 서비스의 접근/이용 | 자기보고 비표준화된 문항 | 정신건강 전문가(심리사, 정신과 전문의 등)로부터 진료 여부 | 호주, 캐나다, 칠레, 콜롬비아, 핀란드, 프랑스, 아일랜드, 일본, 한국, 룩셈부르크, 뉴질랜드, 슬로베니아, 미국 |
| | | 약물 처방 또는 복용 여부(항우울제, 항불안제) | 벨기에, 캐나다, 칠레, 핀란드, 프랑스, 독일, 노르웨이, 슬로베니아, 스페인 |
| 아동 및 청소년의 정신건강 | 표준화된 선별 도구, 진단 및 경험한 증상 | 강점·어려움 설문지(SDQ), KIDSCREEN-27 및 KIDSCREEN-10, 아동 청소년 불안 및 관련 정서장애 선별 척도(SCARED), 단축형 기분 및 감정 설문지(SMFQ) | 호주, 벨기에, 핀란드, 프랑스, 독일, 이탈리아, 슬로베니아, 스페인, 영국, 미국 |
| | | 品行장애, 행동 및 정서적 문제, 긍정적 및 부정적 아동기 경험, 약물 사용/남용 행동에 대한 진단 및 경험 | 캐나다, 이탈리아, 스페인, 튀르키예, 미국 |
| 외로움 및 스트레스 | 표준화된 선별 도구, 비표준화된 자기보고 지표 | 외로움 및 사회적 관계: UCLA 외로움 척도(UCLA Loneliness Scale), 오슬로 사회적 지원 척도(Oslo Social Support Scale), 다차원 사회적 지지 척도(Multidimensional Scale of Perceived Social Support), 비표준화된 지표(non-standardised indicators) | 호주, 오스트리아 , 벨기에, 캐나다, 콜롬비아, 코스타리카, 체코 , 덴마크, 핀란드, 프랑스, 독일, 그리스, 헝가리, 아이슬란드, 아일랜드, 이스라엘, 이탈리아, 일본, 라트비아, 리투아니아, 룩셈부르크, 네덜란드, 뉴질랜드, 노르웨이, 폴란드, 포르투갈, 슬로바키아, 슬로베니아, 스페인, 스웨덴, 미국, 영국 |
| | | 스트레스: 코헨 지각된 스트레스 척도(Cohen Perceived Stress Scale, PSS), 비표준화된 지표(non-standardised indicators) | 캐나다, 콜롬비아, 아이슬란드, 이스라엘, 이탈리아, 한국, 라트비아, 슬로베니아, 스웨덴 |

| 주제 영역 | 도구 유형 | 지표 유형 | 수집 국가 |
|---------------------|---------------------------|--|---|
| 회복탄력성, 낙관주의, 자기 효능감 | 표준화된 복합 척도, 비표준화된 자기보고 지표 | 펠린과 스쿨러의 숙달 척도(Pearlin and Schooler's Mastery Scale), 일반적 자기효능감 척도(General Self-Efficacy Scale), 단축형 회복탄력성 대처 척도(Brief Resilient Coping Scale), 단축형 통합성 설문지(Short Sense of Coherence Questionnaire), 코너-데이비드슨 회복탄력성 척도(Connor-Davidson Resilience Scale, CD RISC-10), 단일 항목 자존감 척도(Single Item Self-esteem Scale), 비표준화된 지표(non-standardised indicators) | 호주, 캐나다, 독일, 이탈리아, 노르웨이, 슬로베니아, 스위스 |
| 정신건강에 대한 태도 | 표준화된 복합 척도, 비표준화된 자기보고 지표 | 낙인, 차별, 정신건강 문제 및 자원에 대한 지식과 이해력 등의 주제를 포괄하는 비표준화된 지표 | 코스타리카, 헝가리, 이탈리아, 일본, 한국, 멕시코, 뉴질랜드, 노르웨이, 슬로베니아, 스웨덴 |
| | | 정신건강 이해능력: 우울증 및 불안 이해능력 설문지(Depression and Anxiety Literacy questionnaire, D-Lit & A-Lit) | 슬로베니아 |

주: 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가에 대한 결과가 표시되어 있다. 영국의 데이터에는 국가 통계청에서 실시한 정신건강 관련 설문조사만 포함되며, 위임된 다른 부서에서 수집된 데이터는 포함되지 않는다. 굵은 글씨로 표시된 국가는 공식 설문조사에서 해당 도구가 보고되지 않았지만, OECD 사무국에서 해당 국가가 오슬로 사회적 지원 척도(OSS-3)가 포함된 제2차, 제3차 유럽건강면접조사(EHIS)에 참여한 것을 근거로 추가되었다.

출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

부속서 2.B. 정신건강을 측정하기 위한 표준화된 조사 도구에 대한 세부 정보

정신건강 문제

정신건강 상태: 구조화된 면담

복합 국제진단면담도구(Composite International Diagnostic Interview, CIDI): 복합 국제진단면담도구(CIDI)는 훈련받은 일반 면담자가 ICD-10 및 DSM-IV의 정의와 기준에 따라 정신장애를 평가하는 데 사용하도록 설계된 포괄적이고 완전하게 구조화된 면담이다(Kessler 및 Bedirhan Üstün, 2006^[26]). 면담의 컴퓨터 지원 버전은 직접 데이터를 입력할 수 있는 소프트웨어 시스템과 함께 사용할 수 있으며, 지필 버전의 면담 응답을 취합하는 데 사용할 수 있다. CIDI는 임상 및 연구 목적뿐만 아니라 역학 및 비교문화 연구 목적에 사용된다. 이를 통해 연구자는 평생 및 12개월 동안의 정신질환 유병률, 해당 질환의 심각도 및 경과, 가정 관리, 직장 생활, 관계 및 사회 생활에 미치는 영향, 서비스 이용 및 약물 사용 등을 측정할 수 있다. 여러 버전의 CIDI가 존재하지만, 가장 최신 버전은 세계보건기구의 복합 국제진단면담도구(WHO-CIDI) V3.0이다(Harvard Medical School, n.d.^[27]). CIDI는 총 40개의 섹션과 선별 모듈로 구성되어 있으며, 이 중 22개 섹션은 진단 섹션으로 기분(2개 섹션), 불안(7개 섹션), 약물 남용(2개 섹션), 아동기(4개 섹션), 기타 장애(7개 섹션)를 평가한다. 나머지 섹션에서는 기능 및 신체적 동반이환, 위험 요인, 사회인구통계학적 정보 및 정신장애의 치료에 대해 평가한다. 선별 모듈은 진단 문항을 심도있게 실시하기 전에 응답자의 일반적인 건강에 대한 기본 문항이 포함되어 있으며, 문항 줄기 응답에 대한 응답자의 피로와 정보 공개를 꺼리는 태도를 줄임으로써 진단 평가의 정확도를 높이는 것으로 나타났다(Harvard Medical School, n.d.^[27]).

정신건강 문제의 증상 선별 도구

로열티가 있는 경우 수수료 및 저작권 제한으로 사용이 어려울 수 있기 때문에, 본 섹션에서는 주로 로열티가 없는 공중보건 도구를 제시한다.

정신건강검사(Mental Health Inventory, MHI-5): 정신건강검사-5(MHI-5)는 심리적 고통의 증상을 선별하기 위한 5가지 항목으로 구성된 척도이다. 이는 38개 항목으로 구성된 정신건강검사(MHI)에서 도출된 것으로, 20개 항목 및 36개 항목으로 구성된 단축형 건강 설문조사(SF-20 및 SF-36)에 포함되어 있다(Berwick 외, 1991^[28], Kelly 외, 2008^[29]). 이 문항은 부정적 정서와 긍정적 정서를 모두 다루며, 세 가지 항목은 기분저하/우울기분에 초점을 맞추고 두 가지 항목은 긴장/불안에 초점을 맞추고 있다(도구 자체는 이러한 측면을 별도로 제시하는 데 사용되지 않음). MHI-5는 정신건강 상태를 측정하는 신뢰할 수 있는 척도로 밝혀졌으며, 다양한 국가의 일반 인구나 환자 표본에서 우울증에 대해 타당화되었고, 더 적은 빈도이긴 하지만 불안장애(범불안 및 공황장애 포함)에 대해서도 타당화되었다(Yamazaki, Fukuhara 및 Green, 2005^[30], Hoeymans 외, 2004^[31], Elovania 외, 2020^[32], Gill 외, 2007^[33], Rumpf 외, 2001^[34], Strand 외, 2003^[35], Thorsen 외, 2013^[36]). 두 가지 불안 관련 항목을 제거해도 MHI가 우울증을 파악하는 효과가 감소하지 않는다는 증거가 있지만, 임상 진단 준거에 따른 공식 진단을 표준 기준으로 사용한 연구에서는 검증되지는 않았다(Yamazaki, Fukuhara 및 Green, 2005^[30]).

표 2.12. MHI-5 설문지(점수표 포함)

| | 항상 | 대부분의 시간 | 상당한 시간 | 가끔 | 약간의 시간 | 전혀 없음 |
|---|----|---------|--------|----|--------|-------|
| 지난 1개월 동안 얼마나 많은 시간 동안 | | | | | | |
| 1. 귀하는 행복한 사람이라고 느꼈습니까? (역코딩) | 1 | 2 | 3 | 4 | 5 | 6 |
| 2. 귀하는 차분함과 편안함을 느꼈습니까? (역코딩) | 1 | 2 | 3 | 4 | 5 | 6 |
| 3. 귀하는 초조함을 느꼈습니까? | 1 | 2 | 3 | 4 | 5 | 6 |
| 4. 귀하는 우울감을 느꼈습니까? | 1 | 2 | 3 | 4 | 5 | 6 |
| 5. 귀하는 기운을 북돋을 수 없을 정도로 우울한 기분을 느꼈습니까? | 1 | 2 | 3 | 4 | 5 | 6 |

주: 각 항목의 점수를 합산하여 최저 5점에서 최대 30점까지의 총점을 산출한 다음, 표준 선형 변환을 통해 0점~100점 범위의 변수로 변환된다. 점수가 높을수록 정신건강이 양호함을 나타내며, 다양한 심리적 고통의 중증도에 대한 준거점수는 다음과 같다. 68점 이하 경증, 중등도 또는 중증, 60점 이하 중등도 또는 중증, 52점 이하 중증을 나타낸다.

출처: Kelly, M.J. 외 (2008_[29]), "Evaluating cutpoints for the MHI-5 and MCS using the GHQ-12: A comparison of five different methods", *BMC Psychiatry* Vol. 8/10, <https://doi.org/10.1186/1471-244X-8-10>.

단축형 건강 설문조사(SF-12): 단축형 건강 설문조사(SF-12)는 건강 관련 삶의 질을 측정하는 도구이다. 일반 인구 및 대규모 설문조사에 사용되는 SF-36 설문지를 축소(단축형)하여 개발되었으며, SF-36의 8개 영역(일반적인 정신건강, 에너지 및 피로, 신체 통증, 건강에 대한 전반적인 인식, 건강으로 인한 신체 활동 제한, 신체적 또는 정서적 상태로 인한 사회 활동 제한, 신체건강으로 인한 일상 활동 제한, 정서적 건강으로 인한 일상 활동 제한)에 대해 최대 2개 항목으로 구성되어 있다(Ware 외, 2002_[37]). SF-12 및 SF-36의 여러 문항은 정신건강검사(MHI)에서 직접 도출된 것으로, 여기에는 MHI-5 독립형 척도도 포함되어 있다(위 참조)(RAND, n.d._[38]). 신체적 요소 요약(PCS) 및 정신건강 요소 요약(MCS)이라는 두 가지 요약 점수가 SF-12에서 도출될 수 있으며, 일반 인구 표본의 최근의 활동성 우울장애에 대해 채점 방법의 범위가 타당화되었고, 더 적은 빈도이기는 하지만 불안장애에 대해서도 타당화되었다 (Ware 외, 2002_[37], Gill 외, 2007_[33], Vilagut 외, 2013_[39]). 일부 증거에 따르면 SF-12의 신체건강 영역과 정신건강 간의 연관성은 저소득 환경에서 더 밀접한 관련이 있을 수 있으며, 상황별 가중치에 영향을 미칠 수 있음을 시사한다(Ohrnberger 외, 2020_[40]). SF-12는 저작권 제한이 적용되므로 본 보고서에 게시할 수 없다(Quality Metric, n.d._[41]).

케슬러 척도(K10/K6): 케슬러 심리적 고통 척도는 주로 10항목(K10) 및 6항목(K6) 형태로 사용되며, 심각한 수준의 심리적 고통을 겪는 성인을 식별하기 위한 선별 도구이다. 이 문항은 신체적 증상과 부정적 정서, 특히 기분저하, 우울기분과 긴장 및 불안에 초점을 맞춘다. 이러한 측면은 일반적으로 별도로 제시되지 않고 고통에 대한 총점이 사용되지만, 요인 분석에 따르면 우울증과 불안은 K10에서 별개의 클러스터로 설정되었다(Brooks, Beard 및 Steel, 2006_[42]). 실제로 일반 인구에 사용하도록 설계되었으나 1차 임상 환경에서도 종종 적용된다. 민감도와 특이도 분석에 따르면 K6과 K10은 지역사회 사례에서 불안과 우울을 식별해 낼 수 있다. (Slade, Grove 및 Burgess, 2011_[43]). 또한 다른 정신장애 범주 및 현재 정신장애의 유무와의 연관성은 낮지만, 케슬러 척도는 여러 문화권에 걸쳐, 우울과 정동장애에 대한 진단적 면담 평가와 비교해 타당화가 진행되었다(Andrews 및 Slade, 2001_[44]). 또한 케슬러 척도가 정신건강 중재가 필요한 중등도 이상의 심리적 고통으로 어려움을 겪고 있는 개인을 포착하는 데 유용하게(준거점수 기준이 낮음) 사용될 수 있다는 증거도 있다(Prochaska 외, 2012_[45]).

표 2.13. 케슬러 척도 10/6 설문지(점수표 포함)

| | 전혀 없음 | 약간의 시간 | 가끔 | 대부분의 시간 | 항상 |
|---|----------|-----------|----|------------|----|
| 귀하는 지난 30일 동안 얼마나 자주 | | | | | |
| 1. 이유없이 피곤하다고 느꼈습니까? | 1 | 2 | 3 | 4 | 5 |
| 2. 초조하다고 느꼈습니까? | 1 | 2 | 3 | 4 | 5 |
| 3. 너무 초조해서 아무것도 귀하를 진정시킬 수 없다고 느꼈습니까? | 1 | 2 | 3 | 4 | 5 |
| 4. 절망스러웠습니까? | 1 | 2 | 3 | 4 | 5 |
| 5. 안전부절못했습니까? | 1 | 2 | 3 | 4 | 5 |
| 6. 가만히 앉아 있을 수 없을 만큼 안전부절못한 적이 있습니까? | 1 | 2 | 3 | 4 | 5 |
| 7. 우울했습니까? | 1 | 2 | 3 | 4 | 5 |
| 8. 피곤했습니까? | 1 | 2 | 3 | 4 | 5 |
| 9. 너무 슬퍼서 아무것도 귀하의 기분을 복돋을 수 없다고 느꼈습니까? | 1 | 2 | 3 | 4 | 5 |
| 10. 무기력하다고 느꼈습니까? | 1 | 2 | 3 | 4 | 5 |

주: 모든 항목의 점수를 합산한 총점이 제공되며, 점수가 높을수록 정신건강이 나쁘다는 것을 나타낸다. 그러나 국가 및 기관에 따라 K10 및 K6 척도에 대해 서로 다른 채점 방법이 사용되었다. 예를 들어 미국은 각 항목이 0점~4점(K10의 경우 최대 총점 40점, K6의 경우 최대 총점 24점)의 척도로 코딩되는 반면, 호주는 상기 표와 같이 1점~5점(K10의 경우 최대 총점 50점, K6의 경우 최대 총점 30점)의 척도로 사용되었다. 호주 건강 설문조사에서 사용되는 K10 점수는 일반적으로 다음과 같다. 5점~10점 심리적 고통이 낮음, 16점~21점 보통, 22점~29점 높음, 30점~50점 매우 높음을 나타낸다. K6 점수의 경우 13점(0점~4점 척도로 코딩) 또는 19점(1점~5점 척도로 코딩) 이상의 점수인 응답자는 일반적으로 심각한 정신질환이 있을 가능성이 있는 것으로 분류된다. 다른 상황에서 준거점수는 다를 수 있다.

출처: ABS(2007_[46]), *Information Paper: Use of the Kessler Psychological Distress Scale in ABS Health Surveys*, Australian Bureau of Statistics, <https://www.abs.gov.au/ausstats/abs@.nsf/lookup/4817.0.55.001chapter92007-08>, Kessler, R. 외 (2010_[47]), "Screening for serious mental illness in the general population with the K6 screening scale: Results from the WHO World Mental Health (WMH) survey initiative", *International Journal of Methods in Psychiatric Research*, Vol. 19/S1, pp. 4-22, <https://doi.org/10.1002/mpr.310>.

일반 건강 설문지(General Health Questionnaire, GHQ-12): 12개 항목으로 구성된 일반 건강 설문지(GHQ-12)는 응답자의 정서(부정적 및 긍정적), 신체적 증상 및 기능 손상에 초점을 맞추어 심리적 고통을 파악하는 척도이다. GHQ-12는 여러 언어로 번역되어 청소년 표본을 포함한 전 세계 일반 및 임상 인구(특히 우울증 및 불안장애에 대해)에서 광범위하게 타당화되었다(Hankins, 2008_[48], Gilbody, 2001_[49], Baksheev 외, 2011_[50]). GHQ-12는 본래 단일 차원성을 검증할 의도였으나 다양한 다차원 구조(예: 불안 및 우울증, 사회적 기능장애, 자신감 상실)를 뒷받침하는 여러 요인 분석 연구와 함께 차원에 대한 논쟁이 있다(Gao 외, 2004_[51]). 그러나 최근 연구에서는 이러한 결과는 항목 문구로 인한 방법론적 변인으로 인해 나타났을 가능성이 높다고 지적하며, 척도를 단일 구성으로 취급하면 편향을 최소화한다는 개념을 뒷받침한다(Hystad 및 Johnsen, 2020_[52]). GHQ-12는 저작권 제한이 적용되므로 본 보고서에 게시할 수 없다.

환자 건강 설문지(PHQ-9/PHQ-8): 전체 환자 건강 설문지(PHQ)에는 기분, 불안, 알코올, 섭식 및 신체형 장애에 중점을 둔 모듈로 구성된 59개 문항이 포함되어 있다. PHQ-9는 우울 증상의 유무와 심각도를 파악하기 위해 설계된 9개의 문항으로 구성된 설문조사이며, 주요우울장애의 DSM-IV 및 DSM-5 증상 기준과 직접 매핑된다. PHQ-8 설문지는 자살 사고에 관한 마지막 문항이 삭제되었다. PHQ-8/9에 대해 단일 요인 구조가 확인되었지만, 최근 연구 결과는 정서적 요인과 신체적 요인으로 구성된 2요인 모델을 뒷받침하고 있다(Sunderland 외, 2019_[53]). 두 도구 모두 다양한 인구 및 임상 환경, 연령대, 문화 및 인종에 걸쳐 수용 가능한 진단 선별 특성을 보여줄 뿐만 아니라 우울증 심각도를 측정하는 신뢰할 수 있고 타당한 척도이기도 하다(Manea, Gilbody 및 McMillan, 2012_[54], Moriarty 외, 2015_[55], Kroenke 외, 2009_[56], Huang 외, 2006_[57], Kroenke, Spitzer 및 Williams, 2001_[58], Richardson 외, 2010_[59]). PHQ-8/9와 DSM 간의 긴밀한

일치성으로 인해 제약이 적은 증상 문항으로 구성된 확장형 자기보고 척도와 비교하여 잠재적으로 서구 중심의 우울증 구조를 포함하여 동일한 비판을 받을 수 있다(Zimmerman 외, 2012^[60], Haroz 외, 2017^[61]).

표 2.14. PHQ-9/8 설문지(점수표 포함)

| | 전혀 없음 | 며칠 | 절반 이상 | 거의 매일 |
|---|-------|----|-------|-------|
| 귀하는 지난 2주 동안 다음과 같은 문제를 얼마나 자주 겪었습니까? | | | | |
| 1. 하는 일에 대한 흥미나 즐거움이 거의 없음 | 0 | 1 | 2 | 3 |
| 2. 울적함, 우울함 또는 절망감 | 0 | 1 | 2 | 3 |
| 3. 잠들기 어렵거나 자주 깬 또는 지나친 수면 | 0 | 1 | 2 | |
| 4. 피곤하거나 기운이 없음 | 0 | 1 | 2 | 3 |
| 5. 식욕부진 또는 과식 | 0 | 1 | 2 | 3 |
| 6. 자신에 대해 부정적으로 느낌 - 자신이 실패자이거나 자신이나 가족을 실망시켰다고 생각함 | 0 | 1 | 2 | 3 |
| 7. 신문 읽기나 텔레비전 시청과 같은 일에 집중하기 어려움 | 0 | 1 | 2 | 3 |
| 8. 다른 사람이 알아차릴 수 있을 정도로 느리게 움직이거나 느리게 말함 또는 그 반대로 안전부절못해서 평소보다 많이 움직임 | 0 | 1 | 2 | 3 |
| 9. 차라리 죽는 편이 낫다고 생각하거나 어떤 식으로든 자신을 해치는 것에 대해 생각 | 0 | 1 | 2 | 3 |

주: 이탤릭체로 표시된 마지막 항목은 자살 충동에 관한 문항으로 PHQ-9에 추가되었다. 채점은 다음과 같이 두 가지 방법으로 할 수 있다. (1) 주요 우울증 또는 기타 우울증에 대한 '진단 알고리즘' 또는 (2) 모든 항목의 점수를 합산하고 우울증 심각도에 대해 서로 다른 준거점수 적용. DSM 정의에 따른 진단 알고리즘에서는 첫 번째 또는 두 번째 항목(우울기분 또는 무쾌감증)이 최소 '절반 이상'으로 나타나고, 전체 증상 중 5개 이상 또는 2개~4개 증상이 최소 '절반 이상'으로 함께 나타나면 각각 주요 우울증 또는 기타 우울증으로 간주한다. 두 번째 형태의 분류에서는 모든 항목의 점수를 합산하여 우울증 심각도의 총점을 산출한다. 점수의 범위는 PHQ-8의 경우 0점~24점, PHQ-9의 경우 0점~27점이다. 0점~4점은 없음, 5점~9점은 경증 우울증, 10점~14점은 중등도 우울증, 15점~19점은 중등도~중증 우울증, 20점~24점/27점은 중증 우울증을 나타낸다. 실제로 10점 이상은 진단 상태와 관계없이 일반적으로 임상적으로 유의미한 우울증을 나타낸다.

출처: Kroenke, K. 외 (2009^[56]), "The PHQ-9 as a measure of current depression in the general population", *Journal of Affective Disorders*, Vol. 114/1-3, pp. 163-173, <https://doi.org/10.1016/j.jad.2008.06.026>, Kroenke, K. 외 (2001^[58]), "The PHQ-9: Validity of a brief depression severity measure", *Journal of General Internal Medicine*, Vol. 16/9, pp. 606-613, <http://dx.doi.org/10.1046/j.1525-1497.2001.016009606.x>.

범불안장애 설문지(GAD-7/GAD-2): 범불안장애 설문지(GAD-7)는 지난 2주 동안 광범위한 불안 관련 문제가 발생한 빈도에 관한 7가지 문항으로 구성되어 있다. 이 척도는 범불안장애의 선별 및 심각도 평가를 위해 개발되었다. 항목은 다수의 범불안장애 증상[DSM-IV 및 5에 명시된 장애의 증상(과도한 걱정, 걱정 통제의 어려움, 안전부절못함 및 과민성 포함, 피로, 근육 긴장, 수면장애는 포함되지 않음)]를 포괄한다. 연구 결과는 척도를 단일 차원 구조로 제시한다(Sunderland 외, 2019^[53]). GAD-7은 환자 및 인구 표본 모두에서 우수한 내적 합치도, 수렴 타당도 및 변화에 대한 민감도를 입증했다(Löwe 외, 2008^[62], Beard 및 Björgvinsson, 2014^[63]). 이 척도는 여러 언어와 현지 방언으로 번역되었으나 도구의 문화 간 편향 가능성에 대한 연구가 앞으로 수행되어야 한다(Parkerson 외, 2015^[64], Sunderland 외, 2019^[53]). 이 척도는 불안의 일반적인 증상에 초점을 맞추고 있으며, 사회불안장애와 같은 다른 불안장애의 유무를 평가하기 위해 개발되지 않았다. 그러나 일부 연구자들은 이 척도가 걱정의 초진단적 과정을 강조하고 범불안장애가 높은 수준의 동반이환율을 보인다는 사실을 고려할 때 다양한 불안장애에 사용될 수 있다고 주장한다(Johnson 외, 2019^[65]). 이 척도의 단축형인 GAD-2는 처음 두 항목(걱정과 걱정 통제의 어려움), 즉 DSM에 따른 범불안의 핵심 기준에만 초점을 맞추고 있다. 여러 연구 결과가 다양한 환경에서 심리측정적 속성과 타당성을

뒷받침하는 것으로 나타났다(Byrd-Bredbenner, Eck 및 Quick, 2021^[66], Hughes 외, 2018^[67], Luo 외, 2019^[68], Ahn, Kim 및 Choi, 2019^[69]).

표 2.15. GAD-7/GAD-2 설문지(점수표 포함)

| | 전혀 없음 | 며칠 | 절반 이상 | 거의 매일 |
|---------------------------------------|-------|----|-------|-------|
| 귀하는 지난 2주 동안 다음과 같은 문제를 얼마나 자주 겪었습니까? | | | | |
| 1. 긴장, 불안 또는 초조한 느낌 | 0 | 1 | 2 | 3 |
| 2. 걱정을 멈출 수 없거나 통제할 수 없음 | 0 | 1 | 2 | 3 |
| 3. 여러 가지 일을 지나치게 걱정함 | 0 | 1 | 2 | 3 |
| 4. 긴장을 푸는 것이 어려움 | 0 | 1 | 2 | 3 |
| 5. 가만히 앉아 있기 힘들 정도로 안절부절못함 | 0 | 1 | 2 | 3 |
| 6. 쉽게 짜증을 내거나 과민해짐 | 0 | 1 | 2 | 3 |
| 7. 끔찍한 일이 일어날 것 같은 두려움을 느낌 | 0 | 1 | 2 | 3 |

주: 이 테이블체로 표시된 항목은 단축형 설문지(GAD-2)에 포함된 2개의 항목을 나타낸다. 모든 항목 점수를 합산하여 총점을 산출하고 점수의 범위는 GAD-7의 경우 0점~21점이며, 점수가 높을수록 불안 증상이 심각함을 나타낸다. 0점~4점은 없음, 5점~9점은 경증 불안, 10점~14점은 중등도 불안, 15점~21점은 중증 불안을 나타낸다. GAD-2의 경우 범불안장애에 대한 진단 평가가 추가적으로 필요한 경우를 식별하는 데 권장되는 준거점수는 3점이다.

출처: Spitzer, R. 외 (2006^[70]), "A brief measure for assessing generalized anxiety disorder: The GAD-7", *Archives of Internal Medicine*, Vol. 166/10, pp. 1092-1097, <http://dx.doi.org/10.1001/ARCHINTE.166.10.1092>.

환자 건강 설문지(PHQ-4): PHQ-4 선별 도구는 우울증과 불안의 핵심 증상 유무와 심각도를 파악하기 위한 단축형 도구로 4개의 문항으로 구성되어 있으며, 일반 인구에서 가장 흔하게 발생하는 두 가지 질환인 우울증과 불안이 함께 동반되는 경우가 많다는 점을 고려한다. PHQ-4의 문항은 PHQ-9/8 및 GAD-7에서 각각 우울증 관련 핵심 문항 2개(이를 합쳐서 PHQ-2라고 함)와 불안 관련 핵심 문항(이를 합쳐서 GAD-2라고 함) 2개를 추출하여 구성되었다. 즉 PHQ-4는 독립적으로 구성 타당도 및 준거 타당도가 입증된 간이형 선별 도구인 PHQ-2와 GAD-2를 결합한 선별 도구이다(위 참조). 일반 인구, 중재, 근로자 및 대학생에 초점을 맞춘 연구에서 PHQ의 심리측정적 속성, 신뢰도 및 타당도가 뒷받침되었다(Stanhope, 2016^[71], Khubchandani 외, 2016^[72], Löwe 외, 2010^[73]).

표 2.16. PHQ-4 설문지(점수표 포함)

| | 전혀 없음 | 며칠 | 절반 이상 | 거의 매일 |
|---------------------------------------|-------|----|-------|-------|
| 귀하는 지난 2주 동안 다음과 같은 문제를 얼마나 자주 겪었습니까? | | | | |
| 1. 긴장, 불안 또는 초조한 느낌 | 0 | 1 | 2 | 3 |
| 2. 걱정을 멈출 수 없거나 통제할 수 없음 | 0 | 1 | 2 | 3 |
| 3. 울적함, 우울함 또는 절망감 | 0 | 1 | 2 | 3 |
| 4. 하는 일에 대해 흥미나 즐거움이 거의 없음 | 0 | 1 | 2 | 3 |

주: 모든 항목 점수를 합산하여 심리적 고통에 대한 총점을 산출하고 점수의 범위는 0점~12점이며, 점수가 높을수록 증상이 심각함을 나타낸다. 0점~2점은 정상, 3점~5점은 경증, 6점~8점은 중등도, 9점~12점은 중증을 나타낸다. 처음 두 항목(GAD-2)의 총점이 3점 이상이면 응답자가 불안 위험군임을 나타낸다. 마지막 두 항목(PHQ-2)의 총점이 3점 이상이면 응답자가 우울증 위험군임을 나타낸다.

출처: Kroenke, K. 외 (2009^[74]), "An ultra-brief screening scale for anxiety and depression: The PHQ-4", *Psychosomatics*, Vol. 50/6, pp. 613-621, <http://dx.doi.org/10.1176/APPL.PSY.50.6.613>.

장애 통계를 위한 워싱턴 그룹의 기능에 대한 간이형 설문지-강화(WG-SS): 워싱턴 그룹의 기능에 대한 간이형 설문지-강화(WG-SS Enhanced)는 국가 통계청, 유엔 기구, 국제 비정부 기구 및 장애인 단체의 대표로

구성된 장애 통계를 위한 워싱턴 그룹에 의해 개발되었고, 그 목적은 인구 및 특별 조사에서 응답자의 장애 유무뿐만 아니라 장애 유형과 중증도를 파악하는 것이다(Washington Group on Disability Statistics, 2020^[75]). 해당 조사는 보기, 듣기, 걷기, 계단 오르기, 기억력 또는 집중력, 자기돌봄, 의사소통, 신체 활동 및 정서 영역의 기능에 중점을 둔다. 후자의 네 가지 문항은 우울증과 불안 증상에 초점을 맞추고 있지만 설문지의 하위 구성요소에 일반적으로 사용되지 않는다. 이와 상관없이 전반적인 기능에 초점을 맞추는 것은 정신 질환의 초진단적 증상을 포착하는 데 중요한 방법을 제공할 수 있다.

표 2.17. WG-SS 강화 설문지

| | 어려움 없음 | 약간 어려움 | 매우 어려움 | 전혀 할 수 없음 | |
|---|--------|--------|--------|-----------|-------|
| 귀하는 다음과 같은 어려움이 있습니까? | | | | | |
| 1. 안경을 착용해도 보는 데 어려움이 있습니까? | | | | | |
| 2. 보청기를 착용해도 듣는 데 어려움이 있습니까? | | | | | |
| 3. 걷거나 계단을 오르는 데 어려움이 있습니까? | | | | | |
| 4. 평소 사용하는 언어로 의사소통 시 이해하는 데 어려움이 있습니까? | | | | | |
| 5. 기억하거나 집중하는 데 어려움이 있습니까? | | | | | |
| 6. 몸을 씻거나 옷을 입는 등 자기돌봄에 어려움이 있습니까? | | | | | |
| 7. 2리터짜리 물병이나 탄산음료를 허리에서 눈높이까지 들어올리는 데 어려움이 있습니까? | | | | | |
| 8. 단추나 연필과 같은 작은 물건을 집거나 용기 또는 병을 여닫는 등 손과 손가락을 사용하는 데 어려움이 있습니까? | | | | | |
| | 매일 | 매주 | 매월 | 일년에 몇 번 | 전혀 없음 |
| 9. 귀하는 걱정, 긴장 또는 불안감을 얼마나 자주 느끼십니까? | | | | | |
| | 조금 | 많이 | 중간 | | |
| 10. 귀하가 마지막으로 느낀 걱정, 긴장 또는 불안감이 어느 정도로 심각합니까? | | | | | |
| | 매일 | 매주 | 매월 | 일년에 몇 번 | 전혀 없음 |
| 11. 귀하는 우울감을 얼마나 자주 느끼십니까? | | | | | |
| | 조금 | 많이 | 중간 | | |
| 12. 귀하가 마지막으로 느낀 우울감이 어느 정도로 심각합니까? | | | | | |

주: 전체 식별자에서 다양한 기능(및 장애의 심각도)별 식별자를 계산할 수 있다. 포함이 권장되는 응답 수준은 다음과 같다. 처음 6개 문항 중 하나 이상에서 ‘매우 어려움’ 또는 ‘전혀 할 수 없음’, 신체 활동과 관련된 2개 문항에 대해 심각도 수준 3 또는 4, 불안 또는 우울증 지표에 대해 심각도 수준 4이 권장된다. 이탤릭체로 표시된 항목은 정신건강 및 신체 기능에 관한 문항을 제외한 6개 항목으로 구성된 간이형 설문지(장애 통계 생산을 위한 워싱턴 그룹의 기능에 대한 간이형 설문지)를 나타낸다.

출처: Washington Group on Disability Statistics (2020^[75]), *The Washington Group Short Set on Functioning: Enhanced (WG-SS Enhanced)*, The Washington Group Data Collection Tools and their Recommended Use (washingtongroup-disability.com).

알코올 사용장애 식별 테스트/단축형(AUDIT/AUDIT-C): 알코올 사용장애 식별테스트(AUDIT)는 1980년대부터 WHO에서 개발한 알코올 선별 검사로 10개 항목으로 구성되어 있다. 이 테스트는 위험 음주자나 알코올 사용장애(알코올 남용 또는 의존 포함) 환자와 응답자를 식별하는 데 도움이 될 수 있다.

입원환자 병동, 응급실, 대학, 직장, 외래 환경 및 정신과 서비스 등 1차 의료기관 이상의 환경에서 타당도가 입증되었다(Berner 외, 2007^[76]). 단축형 버전은 일상적인 환자 면담에 통합할 수 있도록 3개 항목으로 설계되었으며, 확장형과 유사한 정확도를 보이는 것으로 나타났다. 이는 주로 1차 의료기관 환경에서 검증되었으며, 온라인을 통해 주류를 구매하는 성인을 포함한 일반 인구 표본에서의 검증이 확대되었다(Bush 외, 1998^[77], Khadjesari 외, 2017^[78]).

표 2.18. AUDIT/AUDIT-C 설문지(점수표 포함)

| | 전혀 없음 | 월 1회 이하 | 월 2회~4회 | 주 2회~3회 | 주 4회 이상 |
|---|----------|-----------------------|-------------------|---------|-------------|
| 1. 귀하는 알코올이 함유된 음료를 얼마나 자주 드십니까? | 0 | 1 | 2 | 3 | 4 |
| | 1회 또는 2회 | 3회~4회 | 5회~6회 | 7회~9회 | 10회 이상 |
| 2. 귀하는 평소 하루에 알코올이 함유된 음료를 표준잔으로 몇 잔 정도 드십니까? | 0 | 1 | 2 | 3 | 4 |
| | 전혀 없음 | 월 1회 미만 | 매월 | 매주 | 매일 또는 거의 매일 |
| 3. 귀하는 한 번에 6잔 이상의 술을 얼마나 자주 드십니까? | 0 | 1 | 2 | 3 | 4 |
| 4. 귀하는 지난 1년 동안 술을 한 번 마시기 시작하면 자제하지 못한 적이 얼마나 자주 있었습니까? | 0 | 1 | 2 | 3 | 4 |
| 5. 귀하는 지난 1년 동안 술로 인해 일상적으로 해야 할 일에 지장이 생긴 적이 얼마나 자주 있었습니까? | 0 | 1 | 2 | 3 | 4 |
| 6. 귀하는 지난 1년 동안 전날 과음 후 아침에 해장술 마신 적이 얼마나 자주 있었습니까? | 0 | 1 | 2 | 3 | 4 |
| 7. 귀하는 지난 1년 동안 음주 후 죄책감이나 자책감을 얼마나 자주 느끼셨습니까? | 0 | 1 | 2 | 3 | 4 |
| 8. 귀하는 지난 1년 동안 음주 후 전날 밤에 무슨 일이 있었는지 기억하지 못했던 적이 얼마나 자주 있었습니까? | 0 | 1 | 2 | 3 | 4 |
| | 아니오 | 예, 그러나 지난 1년 동안 없었습니다 | 예, 지난 1년 동안 있었습니다 | | |
| 9. 귀하는 음주로 인해 자신이나 타인을 다치게 한 적이 있습니까? | 0 | 2 | 4 | | |
| 10. 친척, 친구, 의사 또는 기타 보건 종사자가 당신의 음주에 대해 걱정하거나 음주를 자제하기를 권한 적이 있습니까? | 0 | 2 | 4 | | |

주: 이탤릭체로 표시된 항목은 단축형 설문지(AUDIT-C)에 포함된 3개의 항목을 나타낸다. 모든 항목의 점수를 합산하여 총점이 산출되며 점수는 0점~40점(AUDIT-C의 경우 0점~12점)의 범위이다. 점수가 높을수록 음주로 인해 안전에 영향을 미칠 가능성이 높다는 것을 나타낸다. AUDIT의 경우 위험하고 유해한 알코올 사용과 알코올 의존성을 나타내는 지표로 준거점수 8점 이상이 권장된다. 알코올은 평균 체중과 신진대사의 차이에 따라 신체에 미치는 영향이 달라지므로 65세 이상인 경우 준거점수를 1점 낮은 7점으로 설정하면 해당 인구집단에 대한 민감도를 높일 수 있다. AUDIT-C의 경우 남성(여성)의 경우 4점(3점) 이상이면 위험한 음주 또는 알코올 사용장애의 증상이 있는 것으로 간주한다.

출처: Bush, K. 외 (1998^[77]), "The AUDIT alcohol consumption questions (AUDIT-C): An effective brief screening test for problem drinking", *Archives of Internal Medicine*, Vol. 158/16, <https://doi.org/10.1001/archinte.158.16.1789>, WHO (2001^[79]), *AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for use in primary health care*, World Health Organization, <https://www.who.int/publications/i/item/audit-the-alcohol-use-disorders-identification-test-guidelines-for-use-in-primary-health-care>.

긍정적 정신건강

OECD 주관적 웰빙 측정 지침의 핵심 문항: OECD 주관적 웰빙 측정 지침에서는 가구 조사에 포함될 수 있는 생애 평가와 (단기적) 정서를 모두 포괄하는 주관적 웰빙에 대한 최소한의 측정을 제안한다(OECD, 2013^[80]). 이에 포함된 핵심 척도는 타당도와 적절성 측면에서 가장 강력한 근거를 가지며, 국제적 비교 가능성이 가장 중요한 척도이다. 행복 웰빙의 한 측면에 대한 실험적 측정도 포함된다.

표 2.19. 주관적 웰빙에 관한 OECD 핵심 문항

| | 0점~10점 |
|---|--------|
| 다음은 삶에서 느끼는 만족감에 관한 문항으로 0점~10점 사이의 척도로 표시합니다. 0점은 '전혀 만족하지 않음', 10점은 '완전히 만족함'을 의미합니다. | |
| 1. 귀하는 전반적으로 요즘 삶에 얼마나 만족합니까? | |
| 다음은 삶에서 하는 일들이 가치 있다고 느끼는 정도에 관한 문항으로 0점~10점의 척도로 표시합니다. 0점은 삶에서 하는 일들이 '전혀 가치가 없음', 10점은 '매우 가치 있음'을 의미합니다. | |
| 2. 귀하는 전반적으로 자신이 하는 일들이 삶에 어느 정도 가치가 있다고 생각합니까? | |
| 다음은 어제 느낀 기분에 관한 문항으로 0점~10점의 척도로 표시합니다. 0점은 어제 '전혀' 경험하지 않았음, 10점은 어제 '항상' 경험했음을 의미합니다. 귀하가 어제 느낀 기분에 대한 문항은 다음과 같습니다. | |
| 3. 귀하는 어제 행복을 어느 정도로 느꼈습니까? | |
| 4. 귀하는 어제 걱정을 어느 정도로 느꼈습니까? | |
| 5. 귀하는 어제 우울을 어느 정도로 느꼈습니까? | |

주: 정서에 관한 세 가지 문항(3번~5번)은 한 묶음으로 포함되어야 하며, 전달 응답자의 정서 상태를 특징짓는 데 필요한 최소한의 문항을 제공하기 위한 것이다.

출처: OECD (2013^[80]), *OECD Guidelines on Measuring Subjective Well-being*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264191655-en>.

WHO-5 웰빙지수(WHO-5): 세계보건기구 정신적 건강상태표(WHO-5)는 응답자의 긍정적인 정서에 초점을 맞춘 5가지 항목으로 간략하게 구성된 설문지이다. 이 설문지는 확장형 WHO/ICD-10 우울증 진단 및 DSM-IV 우울증 척도에서 긍정적 정서의 하위 항목을 선택하여 수정한 것으로, 1998년 WHO 유럽지역 사무소에서 웰빙 측정 프로젝트를 위해 1차 의료기관에 처음 사용한 이후 30개 이상의 언어로 번역되었다(세계보건기구, 1998^[81], Topp 외, 2015^[7]). WHO-5는 임상 시험의 결과 측정뿐만 아니라 우울증에 대한 민감한 선별 도구로서 광범위한 연구 분야와 국가에서 웰빙의 일반적인 척도로 적용되어 왔다(Topp 외, 2015^[7]). 청년층과 노년층을 대상으로 한 연구에 따르면 해당 척도에 대해 단일 차원 구조를 확인했다(Topp 외, 2015^[7]).

표 2.20. WHO-5 설문지(점수표 포함)

| | 항상 | 대부분 | 절반 이상 | 절반 미만 | 가끔 | 전혀 없음 |
|-----------------------------|----|-----|-------|-------|----|-------|
| 지난 2주 동안 | | | | | | |
| 1. 나는 활기가 넘치고 기분이 좋았다 | 5 | 4 | 3 | 2 | 1 | 0 |
| 2. 나는 차분하고 편안함을 느꼈다 | 5 | 4 | 3 | 2 | 1 | 0 |
| 3. 나는 활동적이고 활기찬 기분을 느꼈다 | 5 | 4 | 3 | 2 | 1 | 0 |
| 4. 나는 상쾌한 기분으로 기상하고 휴식을 취했다 | 5 | 4 | 3 | 2 | 1 | 0 |
| 5. 나의 일상엔 흥미로운 일들로 채워졌다 | 5 | 4 | 3 | 2 | 1 | 0 |

주: 모든 항목의 점수를 합산하여 0점~25점 범위로 산출된 총점에 4를 곱하여 총점을 0점(최악의 웰빙)~100점(최상의 웰빙)으로 정규화한다. 50점 이하 또는 52점 이하의 준거점수(Sándor 외, 2021^[82])는 웰빙 저하를 나타내는 지표로 사용되는 경우가 많으며, 이는 우울증 선별과 환자 사망률 예측을 위해 WHO-5를 사용한 연구에서 타당화되었다.

출처: Topp, C. 외 (2015^[7]), “The WHO-5 well-being index: A systematic review of the literature”, *Psychotherapy and Psychosomatics*, Vol. 84/3, pp. 167-176, <https://doi.org/10.1159/000376585>.

SF-36 에너지/활력 하위 척도: SF-36 척도에서 활력에 관한 4가지 항목의 하위 척도(위 참조)는 에너지/피로를 측정하는 일반적인 척도이다. 이 척도는 임상 환경에서 타당화되었으며 확장형 척도(예: 암 관련 피로)에 비해 우수한 성능을 보였다(Brown 외, 2011^[83]).

표 2.21. SF-36 활력 하위 척도 설문지(점수표 포함)

| | 항상 | 대부분 | 상당한 시간 | 가끔 | 약간의 시간 | 전혀 없음 |
|---------------------------------------|----|-----|--------|----|--------|-------|
| 귀하는 지난 4주 동안 다음과 같은 기분을 얼마나 자주 느꼈습니까? | | | | | | |
| 1. 활기가 넘친다고 느꼈습니까? | 1 | 2 | 3 | 4 | 5 | 6 |
| 2. 에너지가 넘친다고 느꼈습니까? | 1 | 2 | 3 | 4 | 5 | 6 |
| 3. 지친 기분이 들었습니까? | 1 | 2 | 3 | 4 | 5 | 6 |
| 4. 피곤함을 느꼈습니까? | 1 | 2 | 3 | 4 | 5 | 6 |

주: 표준화된 점수는 0점~100점의 범위이며 점수가 낮을수록 피로가 심하다는 것을 나타낸다. 45점 이하는 임상적으로 유의한 피로를 나타내는 것으로 확인되었다.

출처: Ware, J. 외 (1993^[84]), *SF-36 Health Survey: Manual and Interpretation Guide*, The Health Institute, New England Medical Center Hospitals, https://www.researchgate.net/profile/John-Ware-6/publication/313050850_SF-36_Health_Survey_Manual_Interpretation_Guide/links/594a5b83aca2723195de5c3d/SF-36-Health-Survey-Manual-Interpretation-Guide.pdf (2023년 1월 22일 접속), Donovan, K. 외 (2008^[85]), “Identifying clinically meaningful fatigue with the Fatigue Symptom Inventory”, *Journal of Pain and Symptom Management*, Vol. 36/5, pp. 480-487, <https://doi.org/10.1016/j.jpainsymman.2007.11.013>.

삶의 만족 척도(SWLS): 삶의 만족 척도는 특정 삶의 영역에 초점을 맞추기보다는 삶의 전반적인 만족도를 평가하기 위해 개발되었다. 초기 연구에 따르면 다른 유형의 주관적 웰빙과 우수한 수렴 타당도를 보이는 동시에 정서적 웰빙 측정과는 구별되는 것으로 나타났다(Pavot 외, 1991^[86], Pavot 및 Diener, 1993^[87]).

표 2.22. SWLS 설문지(점수표 포함)

| | 매우 동의함 | 동의함 | 약간 동의함 | 결정하지 않음 | 약간 동의하지 않음 | 동의하지 않음 | 매우 동의하지 않음 |
|--|--------|-----|--------|---------|------------|---------|------------|
| 1. 내 삶의 대부분이 내 이상에 가깝다 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
| 2. 내 삶의 조건이 훌륭하다 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
| 3. 나는 내 삶에 만족한다. | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
| 4. 지금까지 나는 삶에서 원하는 가장 중요한 것들을 이루었다. | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
| 5. 만약 인생을 다시 살 수 있다면 나는 거의 아무것도 바꾸지 않을 것이다 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |

주: 모든 항목의 점수를 합산하여 총점이 5점~35점 범위로 산출되며 점수가 높을수록 삶의 만족도가 높음을 나타낸다. 5점~9점 매우 불만족, 10점~14점 불만족, 15점~19점 약간 불만족, 20점~24점 약간 만족, 25점~29점 만족, 30점~35점 매우 만족을 나타낸다.

출처: Diener, E. 외 (1985^[88]), “The Satisfaction with Life Scale”, *Journal of Personality Assessment*, Vol. 49/1, pp. 71-75, https://doi.org/10.1207/s15327752jpa4901_13.

단축형 정신건강 연속체 척도(MHC-SF): 단축형 정신건강 연속성 척도(MHC-SF)는 Keyes의 이중연속체 모델에서 긍정적 정신건강을 포착하기 위해 개발한 14개 항목으로 구성된 척도이다(Keyes, 2002^[89]).

이 척도의 항목은 40개 항목으로 구성된 확장형 정신건강 연속체 척도(MHC-LF)에서 추출되었으며 ‘정서적 웰빙’ 항목(정서적 웰빙과 삶의 만족도를 반영) 3개, ‘사회적 웰빙’ 항목 5개, ‘심리적 웰빙’ 항목 (결합될 경우 행복의 웰빙을 반영) 6개로 구성된 별도의 하위 척도로 구성된다(Lamers 외, 2011^[90]). 연구에 따르면 단축형 정신건강 연속체 척도(MHC-SF)에 대해 높은 내적 신뢰도와 중간 수준의 검사-재검사 신뢰도가 나타났고 하위 척도에서 3요인 구조가 확인되었으며, 이는 웰빙 및 기능 측면과 수렴 타당도를 보여준다 (Lamers 외, 2011^[90]).

표 2.23. MHC-SF 설문지(점수표 포함)

| | 전혀 없음 | 1회~2회 | 주 1회 | 주 2회~3회 | 거의 매일 | 매일 |
|--|-------|-------|------|---------|-------|----|
| 귀하는 지난 1개월 동안 다음과 같은 기분을 얼마나 자주 느꼈습니까? | | | | | | |
| <i>정서적 웰빙(정서)</i> | | | | | | |
| 1. 행복 | 0 | 1 | 2 | 3 | 4 | 5 |
| 2. 삶의 관심 | 0 | 1 | 2 | 3 | 4 | 5 |
| 3. 삶의 만족도 | 0 | 1 | 2 | 3 | 4 | 5 |
| <i>사회적 웰빙(행복)</i> | | | | | | |
| 4. 자신이 사회에 기여할 수 있다고 느끼는 것 (사회 공헌) | 0 | 1 | 2 | 3 | 4 | 5 |
| 5. 자신이 공동체(예: 사회 단체, 이웃, 지역, 학교)에 소속되어 있다고 느끼는 것 (사회 통합) | 0 | 1 | 2 | 3 | 4 | 5 |
| 6. 사회가 더 나은 방향으로 나아가고 있다는 것 (사회적 성장) | 0 | 1 | 2 | 3 | 4 | 5 |
| 7. 사람들은 기본적으로 선하다는 것 (사회적 수용) | 0 | 1 | 2 | 3 | 4 | 5 |
| 8. 사회가 작동하는 방식을 이해하는 정도 (사회적 일관성) | 0 | 1 | 2 | 3 | 4 | 5 |
| <i>심리적 웰빙(행복)</i> | | | | | | |
| 9. 자신의 성격이 마음에 드는 정도 (자기 수용) | 0 | 1 | 2 | 3 | 4 | 5 |
| 10. 일상 생활에서의 책임을 관리하는 정도 (환경의 통제) | 0 | 1 | 2 | 3 | 4 | 5 |
| 11. 타인과 따뜻하고 신뢰감 있는 관계를 맺는 정도 (타인과의 긍정적 관계) | 0 | 1 | 2 | 3 | 4 | 5 |
| 12. 자신이 성장하고 더 나은 사람이 되기 위해 도전한 경험 (개인적 성장) | 0 | 1 | 2 | 3 | 4 | 5 |
| 13. 자신의 생각과 의견을 생각하거나 표현하는 자신감 정도 (자율성) | 0 | 1 | 2 | 3 | 4 | 5 |
| 14. 자신의 삶이 방향성이나 의미가 있다고 느끼는 정도 (삶의 목적) | 0 | 1 | 2 | 3 | 4 | 5 |

주: 모든 항목의 점수를 합산하여 0점~ 70점 범위로 총점을 산출하며 점수가 높을수록 긍정적 정신건강이 높은 수준임을 나타낸다. 하위 척도 점수는 정서적 웰빙의 경우 0점~15점, 사회적 웰빙의 경우 0점~25점, 심리적 웰빙의 경우 0점~30점 범위이다. ‘번영(Flourishing)’은 정서적 징후 3개 중 1개 이상, 행복 징후 11개 중 6개 이상(사회적 및 심리적 하위 척도 결합)을 ‘매일’ 또는 ‘거의 매일’ 경험하는 것으로 정의된다. ‘쇠약함(Languishing)’은 정서적 징후 3개 중 1개 이상, 자아 실현적 징후 11개 중 6개 이상을 ‘전혀’ 경험하지 않거나 ‘1회~2회’ 경험한 것으로 정의된다. 번영감이나 쇠약함을 경험하지 않는 개인은 ‘중간 수준의 정신건강’으로 분류된다.

출처: Lamers, S. 외, (2011^[90]), “Evaluating the psychometric properties of the Mental Health Continuum-Short Form (MHC-SF)”, *Journal of Clinical Psychology*, Vol. 67/1, pp. 99-110, <https://doi.org/10.1002/jclp.2074>.

워릭-에딘버러 웰빙지수(WEMWBS): 워릭-에딘버러 웰빙지수(WEMWBS)는 2005년 스코틀랜드 국민건강 보험공단(NHS Health Scotland)의 지원으로 개발되었으며 14개 항목으로 구성되어 있다. 만족감 검사지

(Affectometer)를 시작으로 정신적 웰빙(‘기분이 좋고 기능이 좋은 것’으로 개념화됨)을 측정하기 위해 개발되었다(Warwick Medical School, 2021^[91]). 일부 연구에서는 WEMWBS의 단일 차원 구조를 확인한 반면, 다른 연구에서는 정서적 웰빙, 심리적 기능 또는 행복, 사회적 관계와 관련된 세 가지 잔차 요인을 확인했다(Shannon 외, 2020^[92], Koushede 외, 2019^[93]). 정서에 대한 항목을 줄이고 7개 항목으로 구성된 단축형 워릭-에딘버러 웰빙지수(SWEMWBS)도 사용할 수 있다(Stewart-Brown 외, 2009^[94]). (단축형) 워릭-에딘버러 웰빙지수는 청소년, 임상 표본 및 소수 민족 표본 등 다양한 인구집단과 여러 하위 집단에서 타당화 되었으며 25개 이상의 언어로 번역되어 노르웨이, 스웨덴, 이탈리아, 네덜란드, 덴마크, 독일, 프랑스 및 스페인 인구에서 타당도가 검증되었다. 정신적 웰빙 증진 및 정신질환의 치료와 예방 프로젝트에서 발생하는 변화에 두 척도 모두 민감하게 반응하는 것으로 나타났다(Koushede 외, 2019^[93]). 두 척도 모두 하위 집단 간의 정신적 웰빙을 구분할 수 있지만 단축형 척도인 SWEMWBS는 WEMWBS보다 성별 차이에 덜 민감한 것으로 나타났다(Koushede 외, 2019^[93], Ng Fat 외, 2017^[95]).

표 2.24. (S)WEMWBS 설문지(점수표 포함)

| | 전혀 없음 | 거의 없음 | 가끔 있음 | 자주 있음 | 항상 있음 |
|-------------------------------|-------|-------|-------|-------|-------|
| 지난 2주 동안 | | | | | |
| 1. 나는 미래에 대해 낙관적으로 느꼈다 | 1 | 2 | 3 | 4 | 5 |
| 2. 나는 가치있다고 느꼈다 | 1 | 2 | 3 | 4 | 5 |
| 3. 나는 편안함을 느꼈다 | 1 | 2 | 3 | 4 | 5 |
| 4. 나는 타인에게 관심이 있었다 | 1 | 2 | 3 | 4 | 5 |
| 5. 나는 에너지에 여유가 있었다 | 1 | 2 | 3 | 4 | 5 |
| 6. 나는 문제를 잘 대처했다 | 1 | 2 | 3 | 4 | 5 |
| 7. 나는 생각이 명료했다 | 1 | 2 | 3 | 4 | 5 |
| 8. 내 자신에 대해 좋은 느낌을 받았다 | 1 | 2 | 3 | 4 | 5 |
| 9. 나는 다른 사람과 친하다고 느꼈다 | 1 | 2 | 3 | 4 | 5 |
| 10. 나는 자신감을 느꼈다 | 1 | 2 | 3 | 4 | 5 |
| 11. 나는 어떤 일에 대해 스스로 결정할 수 있었다 | 1 | 2 | 3 | 4 | 5 |
| 12. 나는 사랑받고 있다고 느꼈다 | 1 | 2 | 3 | 4 | 5 |
| 13. 나는 새로운 것에 관심이 생겼다 | 1 | 2 | 3 | 4 | 5 |
| 14. 나는 활기가 넘친다고 느꼈다 | 1 | 2 | 3 | 4 | 5 |

주: 이탤릭체로 표시된 항목은 단축형 설문지(SWEMWBS)에 포함된 7항목을 나타낸다. 14개 항목으로 구성된 척도의 경우 모든 항목의 점수가 합산되어 14점~70점 범위로 총점이 산출된다. 7개 항목으로 구성된 척도의 경우 원점수가 7점~35점의 매트릭 점수로 변환된다(변환표 참조: https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/using/howto/swemwbs_raw_score_to_metric_score_conversion_table.pdf). 두 척도 모두 점수가 높을수록 긍정적 정신건강 수준이 높음을 나타낸다. (S)WEMWBS 점수는 정규 분포에 가깝기 때문에 파라메트릭 분석(parametric analysis)이 가능하다. 범주형 점수의 경우 다음의 두 가지 접근 방식을 사용하여 정신적 웰빙 수준(높음, 평균, 낮음)에 대한 준거점수를 산출할 수 있다. (1) 준거 점수를 +/- 1 표준편차로 설정하여 표본의 약 15%를 높은 수준의 웰빙 범주에, 15%를 낮은 수준의 웰빙 범주에 배치하는 통계적 접근 방식, (2) 검증된 우울증 척도에 대한 벤치마킹 접근 방식(예: 41점~44점은 잠재적/경증 우울증을 나타내고 41점 초과는 임상적 우울증 가능성을 나타냄)으로 역학연구 우울 척도(Center for Epidemiologic Studies Depression Scale, CES-D)를 벤치마크로 사용한다. WEMWBS는 저작권의 보호를 받는다. WEMWBS를 사용하고자 하는 경우 라이선스를 취득할 수 있다. 필요한 라이선스 유형과 신청 방법에 대한 자세한 내용은 <https://warwick.ac.uk/wemwbs/using>에서 확인한다. 공공 부문 조직, 필선 단체, 등록된 사회적 기업, 고등 교육기관에 고용된 연구자에게는 ‘비상업적’ 용도로 사용할 수 있는 무료 라이선스가 제공된다. 자세한 문의 사항은 wemwbs@warwick.ac.uk으로 문의한다.

출처: Warwick-Edinburgh Mental Well-being Scale (WEMWBS) © NHS Health Scotland, University of Warwick and University of Edinburgh, 2006, all rights reserved, Warwick Medical School (2021^[91]), *The Warwick-Edinburgh Mental Wellbeing Scales (WEMWBS)*, <https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/>.

참고문헌

- ABS (2007), *Information Paper: Use of the Kessler Psychological Distress Scale in ABS Health Surveys*, Australian Bureau of Statistics, [46]
<https://www.abs.gov.au/ausstats/abs@.nsf/lookup/4817.0.55.001chapter92007-08>.
- Ahn, J., Y. Kim and K. Choi (2019), “The psychometric properties and clinical utility of the Korean version of GAD-7 and GAD- ”, *Frontiers in Psychiatry*, Vol. 12/10, [69]
<https://doi.org/10.3389/fpsyt.2019.00127>.
- Andrews, G. and T. Slade (2001), “Interpreting scores on the Kessler Psychological Distress Scale (K10)”, *Australian and New Zealand Journal of Public Health*, Vol. 25/6, pp. 494-497, [44]
<https://doi.org/10.1111/j.1467-842X.2001.tb00310.x>.
- Baksheev, G. et al. (2011), “Validity of the 1 -item General Health Questionnaire (GHQ-12) in detecting depressive and anxiety disorders among high school students”, *Psychiatry Research*, Vol. 187/1-2, pp. 291-296, [50]
<https://doi.org/10.1016/j.psychres.2010.10.010>.
- Beard, C. and T. Björgvinsson (2014), “Beyond generalized anxiety disorder: Psychometric properties of the GAD- in a heterogeneous psychiatric sample”, *Journal of Anxiety Disorders*, Vol. 28/6, pp. 547-552, [63]
<https://doi.org/10.1016/j.janxdis.2014.06.002>.
- Berner, M. et al. (2007), “The Alcohol Use Disorders Identification Test for Detecting At-Risk Drinking: A Systematic Review and Meta-Analysis”, *Journal of Studies on Alcohol and Drugs*, Vol. 68/3, pp. 461-473, [76]
<https://doi.org/10.15288/jsad.2007.68.461>.
- Berwick, D. et al. (1991), “Performance of a five-item mental health screening test”, *Medical Care*, Vol. 29/2, pp. 169-176, [28]
<https://doi.org/10.1097/00005650-199102000-00008>.
- Brooks, R., J. Beard and Z. Steel (2006), “Factor structure and interpretation of the K10”, *Psychological Assessment*, Vol. 18/1, pp. 62-70, [42]
<https://doi.org/10.1037/1040-3590.18.1.62>.
- Brown, L. et al. (2011), “Comparison of SF-36 vitality scale and Fatigue Symptom Inventory in assessing cancer-related fatigue”, *Supportive Care in Cancer*, Vol. 19/8, pp. 1255-1259, [83]
<https://doi.org/10.1007/s00520-011-1148-2>.
- Burger, H. and J. Neeleman (2007), “A glossary on psychiatric epidemiology”, *Journal of Epidemiology and Community Health*, Vol. 61/3, pp. 185-189, [4]
<https://doi.org/10.1136/jech.2003.019430>.
- Bush, K. et al. (1998), “The AUDIT alcohol consumption questions (AUDIT-C): An effective brief screening test for problem drinking”, *Archives of Internal Medicine*, Vol. 158/16, [77]
<https://doi.org/10.1001/archinte.158.16.1789>.
- Byrd-Bredbenner, C., K. Eck and V. Quick (2021), “GAD-7, GAD-2, and GAD-mini: Psychometric properties and norms of university students in the United States”, *General Hospital Psychiatry*, Vol. 69, pp. 61-66, [66]
<https://doi.org/10.1016/j.genhosppsych.2021.01.002>.
- Case, A. and A. Deaton (2017), *Mortality and Morbidity in the 21st Century*, Brookings Papers on Economic Activity, [2]
<http://brookings.edu/wp-content/uploads/2017/08/casetextsp17bpea.pdf>.
- Chile, G. (2021), *Saludable Mente [Healthy Mind]*, <https://www.gob.cl/saludablemente/>. [21]

- Diener, E. et al. (1985), “The Satisfaction with Life Scale”, *Journal of Personality Assessment*, Vol. 49/1, pp. 71-75, https://doi.org/10.1207/s15327752jpa4901_13. [88]
- Dobson, K. et al. (2020), “Trends in the prevalence of depression and anxiety disorders among Canadian working-age adults between 000 and 01 ”, *Health Reports*, Vol. 31/12, pp. 12- 23, <https://doi.org/10.25318/82-003-X202001200002-ENG>. [9]
- Donovan, K. et al. (2008), “Identifying clinically meaningful fatigue with the fatigue symptom inventory”, *Journal of Pain and Symptom Management*, Vol. 36/5, pp. 480-487, <https://doi.org/10.1016/j.jpainsymman.2007.11.013>. [85]
- Elovainio, M. et al. (2020), “General Health Questionnaire (GHQ-12), Beck Depression Inventory (BDI-6), and Mental Health Index (MHI-5): Psychometric and predictive properties in a Finnish population-based sample”, *Psychiatry Research*, Vol. 289, p. 112973, <https://doi.org/10.1016/j.psychres.2020.112973>. [32]
- Gao, F. et al. (2004), “Does the 1 -item General Health Questionnaire contain multiple factors and do we need them?”, *Health and Quality of Life Outcomes*, Vol. 2/1, p. 63, <https://doi.org/10.1186/1477-7525-2-63>. [51]
- Gilbody, S. (2001), “Routinely administered questionnaires for depression and anxiety: systematic review”, *BMJ*, Vol. 322/7283, pp. 406-409, <https://doi.org/10.1136/bmj.322.7283.406>. [49]
- Gill, S. et al. (2007), “Validity of the mental health component scale of the 12-item Short-Form Health Survey (MCS-12) as measure of common mental disorders in the general population”, *Psychiatry Research*, Vol. 152/1, pp. 63-71, <https://doi.org/10.1016/j.psychres.2006.11.005>. [33]
- Hankins, M. (2008), “The reliability of the twelve-item general health questionnaire (GHQ-12) under realistic assumptions”, *BMC Public Health*, Vol. 8/1, p. 355, <https://doi.org/10.1186/1471-2458-8-355>. [48]
- Haroz, E. et al. (2017), “How is depression experienced around the world? A systematic review of qualitative literature”, *Social Science & Medicine*, Vol. 183, pp. 151-162, <https://doi.org/10.1016/j.socscimed.2016.12.030>. [61]
- Harvard Medical School (n.d.), *About the WHO WMH-CIDI*, <https://www.hcp.med.harvard.edu/wmhcid/about-the-who-wmh-cidi/>(accessed on 31 May 2022). [27]
- HHS (2021), *U.S. Surgeon General Issues Advisory on Youth Mental Health Crisis Further Exposed by COVID-19 Pandemic*, United States Department of Health and Human Services (HHS), <https://www.hhs.gov/about/news/2021/12/07/us-surgeon-general-issues-advisory-on-youth-mental-health-crisis-further-exposed-by-covid-19-pandemic.html>. [20]
- Hoeymans, N. et al. (2004), “Measuring mental health of the Dutch population: A comparison of the GHQ-12 and the MHI-5”, *Health and Quality of Life Outcomes*, Vol. 2/23, <https://doi.org/10.1186/1477-7525-2-2>. [31]
- Huang, F. et al. (2006), “Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients”, *Journal of General Internal Medicine*, Vol. 21/6, pp. 547-552, <https://doi.org/10.1111/j.1525-1497.2006.00409.x>. [57]

- Hughes, A. et al. (2018), “Diagnostic and clinical utility of the GAD-2 for screening anxiety symptoms in individuals with multiple sclerosis”, *Archives of Physical Medicine and Rehabilitation*, Vol. 99/10, pp. 2045-2049, <https://doi.org/10.1016/j.apmr.2018.05.029>. [67]
- Hystad, S. and B. Johnsen (2020), “The Dimensionality of the 12-Item General Health Questionnaire (GHQ-12): Comparisons of Factor Structures and Invariance Across Samples and Time”, *Frontiers in Psychology*, Vol. 11, <https://doi.org/10.3389/fpsyg.2020.01300>. [52]
- Johnson, S. et al. (2019), “Psychometric properties of the General Anxiety Disorder -Item (GAD-7) scale in a heterogeneous psychiatric sample”, *Frontiers in Psychology*, Vol. 10, <https://doi.org/10.3389/fpsyg.2019.01713>. [65]
- Kelly, M. et al. (2008), “Evaluating cutpoints for the MHI-5 and MCS using the GHQ-12: A comparison of five different methods”, *BMC Psychiatry*, Vol. 8/10, <https://doi.org/10.1186/1471-244X-8-10>. [29]
- Kessler, R. and T. Bedirhan Üstün (2006), “The World Mental Health (WMH) survey initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI)”, *International Journal of Methods in Psychiatric Research*, Vol. 13/2, pp. 93-121, <https://doi.org/10.1002/mpr.168>. [26]
- Kessler, R. et al. (2010), “Screening for serious mental illness in the general population with the K screening scale: results from the WHO World Mental Health (WMH) survey initiative”, *International Journal of Methods in Psychiatric Research*, Vol. 19/S1, pp. 4-22, <https://doi.org/10.1002/mpr.310>. [47]
- Keyes, C. (2002), “The mental health continuum: From languishing to flourishing in life”, *Journal of Health and Social Behavior*, Vol. 43/2, pp. 207-222, <https://doi.org/10.2307/3090197>. [89]
- Khadjesari, Z. et al. (2017), “Validation of the AUDIT-C in adults seeking help with their drinking online”, *Addiction Science & Clinical Practice*, Vol. 12/1, p. 2, <https://doi.org/10.1186/s13722-016-0066-5>. [78]
- Khubchandani, J. et al. (2016), “The psychometric properties of PHQ-4 depression and anxiety screening scale among college students”, *Archives of Psychiatric Nursing*, Vol. 30/4, pp. 457-462, <https://doi.org/10.1016/j.apnu.2016.01.014>. [72]
- KOSIS (n.d.), *Annual prevalence of mental disorders (adjusted for sex and age) (database)*, Korean Statistical Information Service, https://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=TX_117_2009_HB027&comm_path=I2 (accessed on 16 May 2022). [10]
- KOSIS (n.d.), *Depressive disorder prevalence (database)*, National Health and Nutrition Survey, Korean Statistical Information Service, https://knhanes.kdca.go.kr/knhanes/sub01/sub01_05.do#none (accessed on 31 August 2022). [11]
- Koushede, V. et al. (2019), “Measuring mental well-being in Denmark: Validation of the original and short version of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS and SWEMWBS) and cross-cultural comparison across four European settings”, *Psychiatry Research*, Vol. 271, pp. 502-509, <https://doi.org/10.1016/j.psychres.2018.12.003>. [93]

- Kroenke, K., R. Spitzer and J. Williams (2001), “The PHQ-9: Validity of a brief depression severity measure”, *Journal of General Internal Medicine*, Vol. 16/9, <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>. [58]
- Kroenke, K. et al. (2009), “An ultra-brief screening scale for anxiety and depression: The PHQ- ”, *Psychosomatics*, Vol. 50/6, pp. 613-621, <https://doi.org/10.1176/APPI.PSY.50.6.613>. [74]
- Kroenke, K. et al. (2009), “The PHQ-8 as a measure of current depression in the general population”, *Journal of Affective Disorders*, Vol. 114/1-3, pp. 163-173, <https://doi.org/10.1016/j.jad.2008.06.026>. [56]
- Lamers, S. et al. (2011), “Evaluating the psychometric properties of the Mental Health Continuum-Short Form (MHC-SF)”, *Journal of Clinical Psychology*, Vol. 67/1, pp. 99-110, <https://doi.org/10.1002/jclp.20741>. [90]
- Löwe, B. et al. (2008), “Validation and standardization of the Generalized Anxiety Disorder screener (GAD-7) in the general population”, *Medical Care*, Vol. 46/3, pp. 266-274, <https://doi.org/10.1097/mlr.0b013e318160d093>. [62]
- Löwe, B. et al. (2010), “A 4-item measure of depression and anxiety: Validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population”, *Journal of Affective Disorders*, Vol. 122/1-2, pp. 86-95, <https://doi.org/10.1016/j.jad.2009.06.019>. [73]
- Luo, Z. et al. (2019), “Adaptation of the two-item generalized anxiety disorder scale (GAD-2) to Chinese rural population: A validation study and meta-analysis”, *General Hospital Psychiatry*, Vol. 60, pp. 50-56, <https://doi.org/10.1016/j.genhosppsych.2019.07.008>. [68]
- Manea, L., S. Gilbody and D. McMillan (2012), “Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): A meta-analysis”, *Canadian Medical Association Journal*, Vol. 184/3, pp. E191-E196, <https://doi.org/10.1503/cmaj.110829>. [54]
- Moriarty, A. et al. (2015), “Screening and case finding for major depressive disorder using the Patient Health Questionnaire (PHQ-9): A meta-analysis”, *General Hospital Psychiatry*, Vol. 37/6, pp. 567-576, <https://doi.org/10.1016/j.genhosppsych.2015.06.012>. [55]
- Mueller, A. and D. Segal (2015), “Structured versus Semistructured versus Unstructured Interviews”, in Cautin, R. and S. Lilienfeld (eds.), *The Encyclopedia of Clinical Psychology*, John Wiley & Sons, Inc, https://www.researchgate.net/profile/Daniel-Segal-6/publication/313966231_Structured_versus_Semistructured_versus_Unstructured_Interviews/links/599a13bcaca272e41d3ec727/Structured-versus-Semistructured-versus-Unstructured-Interviews.pdf (accessed on 30 March 2022). [3]
- National Academies of Sciences Engineering and Medicine (2021), *Mental Health, Substance Use, and Wellbeing in Higher Education: Supporting the Whole Student*, The National Academies Press, Washington, D.C., <https://doi.org/10.17226/26015>. [6]
- National Center for Health Statistics (2021), *Estimates of Mental Health Symptomatology, by Month of Interview: United States, 2019 (database)*, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, <https://www.cdc.gov/nchs/data/nhis/mental-health-monthly-508.pdf> (accessed on 16 May 2022). [13]

- Ng Fat, L. et al. (2017), “Evaluating and establishing national norms for mental wellbeing using the short WarwickEdinburgh Mental Well-being Scale (SWEMWBS): Findings from the Health Survey for England”, *Quality of Life Research*, Vol. 26/5, pp. 1129-1144, <https://doi.org/10.1007/s11136-016-1454-8>. [95]
- OECD (2021), *A New Benchmark for Mental Health Systems: Tackling the Social and Economic Costs of Mental Ill-Health*, OECD Health Policy Studies, OECD Publishing, Paris, <https://doi.org/10.1787/4ed890f6-en>. [1]
- OECD (2021), *COVID-19 and Well-being: Life in the Pandemic*, OECD Publishing, Paris, <https://doi.org/10.1787/1e1ecb53-en>. [18]
- OECD (2021), *Health at a Glance*, OECD Publishing, Paris, <https://doi.org/10.1787/19991312>. [17]
- OECD (2021), *Supporting young people’s mental health through the COVID-19 crisis*, OECD Policy Responses to Coronavirus (COVID-19), OECD Publishing, Paris, <https://www.oecd.org/coronavirus/policy-responses/supporting-young-people-s-mental-health-through-the-covid-19-crisis-84e143e5/>. [19]
- OECD (2020), *How’s Life? 2020: Measuring Well-being*, OECD Publishing, Paris, <https://doi.org/10.1787/9870c393-en>. [16]
- OECD (2013), *OECD Guidelines on Measuring Subjective Well-being*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264191655-en>. [80]
- OECD/European Union (2018), *Health at a Glance: Europe 2018: State of Health in the EU Cycle*, OECD Publishing, Paris/European Union, Brussels, https://doi.org/10.1787/health_glance_eur-2018-en. [23]
- Ohrnberger, J. et al. (2020), “Validation of the SF1 mental and physical health measure for the population from a low-income country in sub-Saharan Africa”, *Health and Quality of Life Outcomes*, Vol. 18/78, <https://doi.org/10.1186/s12955-020-01323-1>. [40]
- Parkerson, H. et al. (2015), “Cultural-based biases of the GAD-7”, *Journal of Anxiety Disorders*, Vol. 31, pp. 38-42, <https://doi.org/10.1016/j.janxdis.2015.01.005>. [64]
- Pavot, W. and E. Diener (1993), “Review of the Satisfaction With Life Scale.”, *Psychological Assessment*, Vol. 5/2, pp. 164-172, <https://doi.org/10.1037/1040-3590.5.2.164>. [87]
- Pavot, W. et al. (1991), “Further validation of the Satisfaction with Life Scale: evidence for the cross-method convergence of well-being measures”, *Journal of Personality Assessment*, Vol. 57/1, pp. 149-161, https://doi.org/10.1207/s15327752jpa5701_17. [86]
- Prochaska, J. et al. (2012), “Validity study of the K scale as a measure of moderate mental distress based on mental health treatment need and utilization”, *International Journal of Methods in Psychiatric Research*, Vol. 21/2, pp. 88-97, <https://doi.org/10.1002/mpr.1349>. [45]
- Public Health Agency Sweden (2022), *Synen på psykisk ohälsa och suicid*, <https://www.folkhalsomyndigheten.se/publicerat-material/publikationsarkiv/s/synen-pa-psykisk-ohalsa-och-suicid-/?pub=105538#105580>. [25]
- Quality Metric (n.d.), *The SF-12v2 PRO Health Survey*, <https://www.qualitymetric.com/health-surveys-old/the-sf-12v2-health-survey/> (accessed on 10 May 2022). [41]

- RAND (n.d.), *Mental Health Inventory Survey*, https://www.rand.org/health-care/surveys_tools/mos/mental-health.html (accessed on 10 May 2022). [38]
- Richardson, L. et al. (2010), “Evaluation of the Patient Health Questionnaire-9 item for detecting major depression among adolescents”, *Pediatrics*, Vol. 126/6, pp. 1117-1123, <https://doi.org/10.1542/peds.2010-0852>. [59]
- Ruedgers, R. (2001), *Handbook of diagnostic and structured interviewing*, Guilford Press, <https://psycnet.apa.org/record/2001-05047-000>. [5]
- Rumpf, H. et al. (2001), “Screening for mental health: validity of the MHI-5 using DSM-IV Axis I psychiatric disorders as gold standard”, *Psychiatry Research*, Vol. 105/3, pp. 243-253, [https://doi.org/10.1016/S0165-1781\(01\)00329-8](https://doi.org/10.1016/S0165-1781(01)00329-8). [34]
- SAMHSA (2019), *Key Substance Use and Mental Health Indicators in the United States: Results from the 2018 National Survey on Drug Use and Health (database)*, Substance Abuse and Mental Health Services Administration, Rockville, MD, <https://www.samhsa.gov/data/sites/default/files/cbhsq-reports/NSDUHNationalFindingsReport2018/NSDUHNationalFindingsReport2018.pdf> (accessed on 16 May 2022). [12]
- Sándor, E. et al. (2021), *Impact of COVID-19 on young people in the EU*, Publications Office of the European Union, <https://www.eurofound.europa.eu/publications/report/2021/impact-of-covid-19-on-young-people-in-the-eu#:~:text=COVID%2D19%20pandemic-,Young%20people%20were%20more%20likely%20than%20older%20groups%20to%20experience,home%20requirements%20and%20school%20closures>. [82]
- Santé Publique France (2021), *La santé mentale au temps de la COVID-19 : En parler, c’est déjà se soigner [Mental health in the time of COVID-19: Talking about it is already treating yourself]*, Santé Publique France, <https://www.santepubliquefrance.fr/presse/2021/la-sante-mentale-au-temps-de-la-covid-19-en-parler-c-est-deja-se-soigner>. [22]
- Santomauro, D. et al. (2021), “Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic”, *The Lancet*, Vol. 398/10312, pp. 1700-1712, [https://doi.org/10.1016/S0140-6736\(21\)02143-7](https://doi.org/10.1016/S0140-6736(21)02143-7). [24]
- Shannon, S. et al. (2020), “Testing the factor structure of the Warwick-Edinburgh Mental Well-Being Scale in adolescents: A bi-factor modelling methodology”, *Psychiatry Research*, Vol. 293, <https://doi.org/10.1016/j.psychres.2020.113393>. [92]
- Slade, T., R. Grove and P. Burgess (2011), “Kessler Psychological Distress Scale: Normative Data from the 00 Australian National Survey of Mental Health and Wellbeing”, *Australian & New Zealand Journal of Psychiatry*, Vol. 45/4, pp. 308-316, <https://doi.org/10.3109/00048674.2010.543653>. [43]
- Spitzer, R. et al. (2006), “A brief Measure for assessing Generalized Anxiety Disorder: The GAD-7”, *Archives of Internal Medicine*, Vol. 166/10, pp. 1092-1097, <https://doi.org/10.1001/ARCHITE.166.10.1092>. [70]
- Stanhope, J. (2016), “Patient Health Questionnaire-4”, *Occupational Medicine*, Vol. 66/9, pp. 760-761, <https://doi.org/10.1093/occmed/kqw165>. [71]

- Statistics Canada (2013), *Canadian Community Health Survey: Mental Health, 2012*, The Daily, [8]
<https://www150.statcan.gc.ca/n1/daily-quotidien/130918/dq130918a-eng.htm> (accessed on 16 May 2022).
- Stewart-Brown, S. et al. (2009), “Internal construct validity of the Warwick-Edinburgh Mental [94]
 Well-being Scale (WEMWBS): a Rasch analysis using data from the Scottish Health Education
 Population Survey”, *Health and Quality of Life Outcomes*, Vol. 7/1, p. 15,
<https://doi.org/10.1186/1477-7525-7-15>.
- Strand, B. et al. (2003), “Measuring the mental health status of the Norwegian population: A [35]
 comparison of the instruments SCL-25, SCL-10, SCL-5 and MHI-5 (SF-36)”, *Nordic Journal of
 Psychiatry*, Vol. 57/2, pp. 113-118, <https://doi.org/10.1080/08039480310000932>.
- Sunderland, M. et al. (2019), “Self-Report Scales for Common Mental Disorders”, in *The Cambridge [53]
 Handbook of Clinical Assessment and Diagnosis*, Cambridge University Press,
<https://doi.org/10.1017/9781108235433.019>.
- Thorsen, S. et al. (2013), “The predictive value of mental health for long-term sickness absence: the [36]
 Major Depression Inventory (MDI) and the Mental Health Inventory (MHI-5) compared”, *BMC
 Medical Research Methodology*, Vol. 13/115, <https://doi.org/10.1186/1471-2288-13-115>.
- Topp, C. et al. (2015), “The WHO-5 well-being index: A systematic review of the literature”, [7]
Psychotherapy and Psychosomatics, Vol. 84/3, pp. 167-176, <https://doi.org/10.1159/000376585>.
- University of Essex, Institute for Social and Economic Research (2022), *Understanding Society: Waves [15]
 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009 (database)*, 5th Edition. UK Data
 Service., <https://doi.org/10.5255/UKDA-SN-6614-16> (accessed on 10 June 2022).
- Vilagut, G. et al. (2013), “The mental component of the Short-Form 12 health survey (SF-12) as a [39]
 measure of depressive disorders in the general population: Results with three alternative scoring
 methods”, *Value in Health*, Vol. 16/4, pp. 564-573, <https://doi.org/10.1016/j.jval.2013.01.006>.
- Ware, J. et al. (1993), *SF-36 Health Survey: Manual and Interpretation Guide*, The Health Institute, [84]
 New England Medical Center Hospitals,
[https://www.researchgate.net/profile/John-Ware-6/publication/313050850_SF-36_Health_Survey_Manual
 Interpretation_Guide/links/594a5b83aca2723195de5c3d/SF-36-
 Health-Survey-Manual-Interpretation-Guide.pdf](https://www.researchgate.net/profile/John-Ware-6/publication/313050850_SF-36_Health_Survey_Manual_Interpretation_Guide/links/594a5b83aca2723195de5c3d/SF-36-Health-Survey-Manual-Interpretation-Guide.pdf) (accessed on 22 January 2023).
- Ware, J. et al. (2002), *How to score SF-12 items: How to score version 2 of the SF-12 Health Survey*, [37]
<https://www.researchgate.net/publication/291994160>.
- Warwick Medical School (2021), *The Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS)*, [91]
<https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/>.
- Washington Group on Disability Statistics (2020), *The Washington Group Short Set on Functioning: [75]
 Enhanced (WG-SS Enhanced)*, [https://www.washingtongroup-
 disability.com/fileadmin/uploads/wg/Documents/Washington_Group_Questionnaire_3_-_WG_Short_Set
 on_Functioning_-_Enhanced.pdf](https://www.washingtongroup-disability.com/fileadmin/uploads/wg/Documents/Washington_Group_Questionnaire_3_-_WG_Short_Set_on_Functioning_-_Enhanced.pdf).

- Woodhead, C. et al. (2012), “Impact of exposure to combat during deployment to Iraq and Afghanistan on mental health by gender”, *Psychological Medicine*, Vol. 42/9, pp. 1985-1996, <https://doi.org/10.1017/S003329171100290X>. [14]
- World Health Organization (2001), *AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for use in primary health care*, World Health Organization, <https://www.who.int/publications/i/item/audit-the-alcohol-use-disorders-identification-test-guidelines-for-use-in-primary-health-care>. [79]
- World Health Organization (1998), *Wellbeing measures in primary health care/the DepCare Project: Report on a WHO meeting*, World Health Organization, <https://apps.who.int/iris/handle/10665/349766>. [81]
- Yamazaki, S., S. Fukuhara and J. Green (2005), “Usefulness of five-item and three-item Mental Health Inventories to screen for depressive symptoms in the general population of Japan”, *Health and Quality of Life Outcomes*, Vol. 3/1, pp. 1-7, <https://doi.org/10.1186/1477-7525-3-48/FIGURES/1>. [30]
- Zimmerman, M. et al. (2012), “How can we use depression severity to guide treatment selection when measures of depression categorize patients differently?”, *Journal of Clinical Psychiatry*, Vol. 73/10, pp. 1287-1291, <https://doi.org/10.4088/JCP.12m07775>. [60]

주

¹ 이는 임상 면담을 통해 도달한 진단의 경우, 그 진단의 근거가 되는 분류 체계에 한하여 유효하다는 것을 의미한다(Mueller 및 Segal, 2015^[3])(3장의 상자 3.4 참조).

² 가구 조사의 범위도 전체 인구를 포괄하지 않으며 표본으로 추출된 군집만 포함된다. 일반적으로 일반 인구보다 정신질환 유병률이 높을 가능성이 높은 노숙자와 시설 환경에 거주하는 사람들은 고려되지 않는다.

³ 설문조사에 응답한 국가는 호주, 오스트리아, 벨기에, 불가리아, 캐나다, 칠레, 콜롬비아, 코스타리카, 체코, 덴마크, 핀란드, 프랑스, 독일, 그리스, 헝가리, 아이슬란드, 아일랜드, 이스라엘, 이탈리아, 일본, 한국, 라트비아, 리투아니아, 룩셈부르크, 멕시코, 네덜란드, 노르웨이, 뉴질랜드, 폴란드, 포르투갈, 슬로베니아, 슬로바키아, 스페인, 스웨덴, 스위스, 튀르키예, 영국 및 미국이다.

⁴ 또한 OECD는 인구 10만 명당 정신과 전문의, 심리사 또는 정신건강 전문가 수, 정신건강 관리 전용 병상수, 정신건강 관련 의료 서비스 지출 등 정신건강 관련 의료 서비스에 대한 행정적 데이터를 발표하고 있다(OECD, 2021^[17]). 이는 인구 수준의 정신건강 *결과*로 간주되지 않으므로 본 프로젝트의 목적에 따라 더 이상 고려하지 않는다.

⁵ 일부 국가에서는 신규 설문조사를 단독으로 실시하고 기존 설문조사에 정신건강 모듈을 추가하는 등 두 가지 방식을 동시에 수행하여 백분율이 68%에 미치지 못한다.

⁶ 또한 Eurostat에서 실시한 2013년 및 2018년 유럽연합 소득 및 생활조건 통계(EU-SILC) 조사의 특별 웰빙 모듈에는 MHI-5가 포함되어 있었으나 그 이후 웰빙 모듈에서 제거되었다는 점에 유의할 필요가 있다. 그 결과 일부 개별 회원국은 국가 건강 및 웰빙 조사에서 MHI-5 척도를 유지하기로 선택할 수 있지만 향후 해당 척도의 사용이 크게 감소할 수 있다.

⁷ 정신질환에 대한 태도 및 낙인 측정에 사용되는 설문조사에 대한 자세한 논의는 OECD(2021^[11])의 표 6.2를 참조한다.

3

가구 조사를 통한 인구 정신건강 측정의 모범 사례

현재 모든 OECD 국가에서 인구 수준의 정신건강을 측정하고 있으나 동일한 구성개념을 측정하는 데 다양한 도구를 사용하고 있어 다종의 결과가 포착되고 있다. 본 장에서는 정신건강 결과의 통일성을 개선하기 위해 적절한 조사 도구를 선택할 때 고려해야 할 기준을 제시한다. 이러한 기준에는 통계 품질, 현장 조사의 실용성, 데이터 분석이 포함된다. 전반적으로 정신건강 문제 및 긍정적 정신건강의 복합 척도에 가장 일반적으로 사용되는 선별 도구의 통계적 적절성을 뒷받침하는 강력한 근거가 있다. 이에 따라 기존의 데이터 수집 관행에 더해 정신건강 스펙트럼 전반의 결과를 포착하는 4가지 도구(PHQ-4, WHO-5 또는 SWEMWBS, 일반적인 정신건강 상태에 관한 문항)를 가구 조사에 포함할 것을 제안하는 바이다.

OECD 국가에서는 이미 다양한 조사 도구를 사용하여 인구 수준의 정신건강 측면을 측정하고 있다. 제2장에서는 우울증 위험, 생애 평가, 일반적인 심리적 고통과 같은 결과에는 어느 정도 통일성이 있으나 불안, 기타 특정 정신장애(양극성장애, 외상 후 스트레스 장애, 섭식장애 등), 긍정적 정신건강의 정서 및 행복 측면 등에서는 조사 범위의 격차가 존재한다는 점을 강조했다. 본 장에서는 구체적인 권장사항을 도출하기 전에 가구 조사에서 정신건강 결과를 측정하는 데 고려해야 할 도구별 특성에 대해 전반적인 개요를 제공한다.

OECD 국가에서는 기존에 구조화된 면담, 이전 진단 데이터, 증상 경험, 자살 사고 및 자살 시도에 대한 문항 등의 다양한 도구를 사용하고 있으나 본 장에서는 선별 도구¹, 긍정적 정신건강 지표, 정신건강 상태에 대한 일반 문항 등 일반적으로 사용되는 세 가지 도구의 통계적 품질에 초점을 맞추고 있으며 이는 크게 두 가지 이유 때문이다. 첫째, 해당 도구는 문항 구성 측면에서 표준화되어 있으므로 통일된 권장사항을 제시하는 데 가장 용이한 기반이 된다. 둘째 해당 도구는 건강 관련 설문조사보다 더 자주 수집되는 경향이 있는 일반 사회조사(진단 또는 경험한 증상에 대한 도구에 비해)에서 일반적으로 사용된다. 세 가지 도구를 종합하면 단일 연속체와 이중 연속체에서 개념화된 결과를 모두 포괄하는 정신건강의 총체적인 측정치를 제공하며(그림 1.2 참조) 자살 사고이나 시도에 대한 측정치보다 더 많은 결과를 제공한다. 따라서 해당 도구는 실용성에 대해 고려할 때 많은 국가에서 채택할 가능성이 가장 높은 도구이다.

적절한 도구 선택에서 가장 중요하게 고려해야 할 사항은 각국 정부가 정신건강에 대한 통합적 정책 접근 방식의 일환으로 해당 도구를 사용할 수 있도록 국가, 집단 및 시간에 걸쳐 정신건강의 다양한 측면을 가장 정확하게 측정할 수 방법에 관한 것이다. 사회적 조건(경제, 사회, 환경)이 다양한 인구집단의 정신건강에 어떻게 영향을 미치는지, 그리고 이러한 조건이 정신건강의 개선 또는 악화에 어떻게 기여하는지 등을 파악하기 위해서는 고품질의 데이터가 필요하다. 측정 편향이 완전히 존재하지 않는 데이터는 없으므로 데이터 수집 주체가 측정의 노이즈 양을 최소화하기 위해 엄격한 품질 관리를 시행하는 것이 매우 중요하다. 그러나 설문조사 응답 행동에 영향을 미치는 낙인과 편견, 다양한 문화적 관점, 시간이 경과함에 따라 변화하는 정신건강에 대한 태도 등으로 인해 정신건강 측정에는 어려움이 있다. 또한 가구 조사는 장기 요양 시설, 병원, 교도소 등 시설에 수용된 인구와 영구 주소가 없는 사람 등 일부 정신건강 질환에 대해 평균보다 높은 위험도를 보일 수 있는 인구집단이 제외되어 있다.

인구 수준에서 정신건강을 측정하는 모범사례는 임상 수준에서 정신건강을 측정하는 모범 사례와는 여러 가지 면에서 상이하다. 국가 통계청이나 보건부에서 전국 인구를 대상으로 대규모 설문조사를 실시하는 경우 임상적 관점에서 장시간 동안 구조화된 면담을 실시하는 것은 표준 기준으로 간주될 수 있다 할지라도, 비현실적이다. 데이터의 최종 사용자는 임상가와 다르며 정책 입안자는 전반적인 추세(시간 경과에 따른 변화, 위험군 간 차이, 국가 간 차이) 추적과 인구 집단의 위험 및 회복탄력성 요인을 추적하는 등, 질환을 진단하고 치료계획을 개발하고자 하는 임상가와 다른 요구사항을 가지고 있다. 본 장은 이러한 필요성을 바탕으로 논의가 진행된다.

본 장에서는 데이터 수집기관에서 고려해야 할 일련의 문항을 제시하면서 고품질의 인구 수준 정신건강 결과 데이터를 생산하는 모범 사례에 대한 지침을 제공한다. 이 사례에서 도출된 주요 결과는 아래 표 3.1에 표시되어 있다. 표에 포함된 특정 선별 도구 및 복합 척도 도구는 OECD 국가에서 가장 일반적으로 사용되는 도구이다(각 도구에 대한 자세한 내용은 표 2.7, 표 2.11, 표 2.12 및 부속서 2.B 참조)². 문항은 (1) 통계 품질, (2) 데이터 수집 절차, (3) 분석 등 크게 세 가지 범주로 분류된다. OECD 국가에서 사용하는 모든 정신건강 도구를 포괄적으로 평가하기보다는 기존 연구에서 도출된 근거를 바탕으로 각 문항의 영역을 설명한다. 이러한 구성을 통해 인구 정신건강을 측정하기 위한 다양한 도구의 장단점을 평가하고 결론에서는 도구의 활용 및 통일화에 대한 구체적인 권장사항을 도출한다.

표 3.1. 통계 품질, 데이터 수집 및 분석 지표에 대한 정신건강 도구 성능 개요

| 도구 정보 | | 통계 품질 | | | | | 데이터 분석 | | 국가 범위 | |
|--|--|--------|-----|-----|------------|---------------------|------------------|-------|-----------------|-------------|
| 도구명 | 주제 범위 및 항목 길이 | 기준 기간 | 신뢰도 | 타당도 | 결측비율 낮음 | 집단 간 비교가능성 높음 | 변화에 대한 민감도 | 정규 분포 | 임계값의 민감도/특이도 | 국가 범위 |
| 정신건강 문제 평가를 위한 검증된 선별 도구 | | | | | | | | | | |
| <i>심리적 고통</i> | | | | | | | | | | |
| 일반 건강 설문지 (GHQ-12) | 부정적 및 긍정적 정서, 신체적 증상, 기능 손상 (12개 항목) | 최근 | ✓ | × | × | ~ | | × | × | 37개국 중 5개국 |
| 케슬러 척도 6(K6) | 부정적 정서 (6개 항목) | 지난 4주 | ○ | ✓ | | ~ | | | ✓ | 37개국 중 4개국 |
| 케슬러 척도 10(K10) | 부정적 정서, 기능 손상 (10개 항목) | 지난 4주 | ✓ | ✓ | | ~ | | | ✓ | 37개국 중 4개국 |
| 정신건강검사-5 (MHI-5) | 부정적 및 긍정적 정서 (5개 항목) | 지난 1개월 | ✓ | ✓ | | | | × | ~ | 37개국 중 28개국 |
| <i>우울 증상</i> | | | | | | | | | | |
| 환자 건강 설문지-8 또는 환자 건강 설문지-9 (PHQ-8/PHQ-9) | 부정적 정서, 무쾌감증, 신체적 증상, 기능 손상(DSM-IV 및 DSM-5 기준에 따라 주요 우울장애와 일치) (8개 또는 9개 항목) | 지난 2주 | ✓ | ✓ | | ✓ | | | ✓ | 37개국 중 30개국 |
| 환자 건강 설문지-2 (PHQ-2) | 부정적 정서, 무쾌감증(2개 항목) | 지난 2주 | ~ | ~ | | ~ | | | ✓ | 37개국 중 8개국 |
| 역학연구센터 우울 척도(CES-D) | 부정적 정서, 무쾌감증(20개 항목) | 지난 1주 | ✓ | ~ | | × | | | × | 37개국 중 17개국 |
| <i>불안 증상</i> | | | | | | | | | | |
| 범불안장애 척도-7 (GAD-7) | 부정적 정서, 신체적 증상, 기능 손상 (7개 항목) | 지난 2주 | ✓ | ✓ | | ~ | ○ | | ~ | 37개국 중 11개국 |

| 도구 정보 | | | | 통계 품질 | | | | 데이터 분석 | | 국가 범위 |
|--------------------------------------|---|----------------------------|-----|-------|------------|---------------------|------------------|--------|-----------------|----------------|
| 도구명 | 주제 범위 및 항목 길이 | 기준 기간 | 신뢰도 | 타당도 | 결측비율 낮음 | 집단 간 비교가능성 높음 | 변화에 대한 민감도 | 정규 분포 | 임계값의 민감도/특이도 | 국가 범위 |
| 범불안장애 척도-2 (GAD-2) | 부정적 정서, 기능 손상 (2개 항목) | 지난 2주 | ✓ | ✓ | ~ | ~ | ~ | ~ | ~ | 37개국 중 7개국 |
| 우울증 및 불안 증상 원자 건강 질문지-4(PHQ-4) | 부정적 정서, 무쾌감증, 기능 손상 (4개 항목) | 지난 2주 | ~ | ~ | ~ | ~ | ~ | ~ | ~ | 37개국 중 13개국 |
| 긍정적 정신건강 평가를 위한 표준화된 도구 | | | | | | | | | | |
| 단축형 건강 상태(SF-12) | 긍정적 및 부정적 정서, 기능 손상(정신건강 요소 요약) (12개 항목) | 지난 4주 | ✓ | ✓ | × | × | | | | 37개국 중 8개국* |
| 윌릭-에딘버러 웰빙지수(WEMWBS) | 긍정적 정서, 행복, 사회적 웰빙 (14개 항목) | 지난 2주 | ✓ | ✓ | ✓ | ~ | ✓ | ✓ | ~ | 37개국 중 27개국 |
| 단축형 윌릭-에딘버러 웰빙지수 (SWEWBS) | 긍정적 정서, 행복, 사회적 웰빙 (7개 항목) | 지난 2주 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ~ | 37개국 중 6개국 |
| WHO-5 웰빙지수(WHO-5) | 긍정적 정서 (5개 항목) | 지난 2주 | ✓ | ✓ | ✓ | ✓ | | | | 37개국 중 6개국 |
| 단축형 정신건강 연속성 척도(MHC-SF) | 긍정적 정서, 행복, 삶의 만족도, 사회적 웰빙 (14개 항목) | 지난 1개월 | ✓ | ~ | ~ | ~ | | | × | 37개국 중 2개국 |
| 일반적인 정신건강 상태에 대한 자기보고(단일 문항) | | | | | | | | | | |
| 자기보고 정신건강(SRMH) | 일반적인 정신건강 상태에 대한 자기보고, 정신적으로 건강한 일수 정신질환으로부터의 회복, 정신건강에 대한 만족도, 정신건강이 일상생활에 지장을 초래하는 정도 등 매우 다양한 (단일 문항) | 다양함 (현재부터 지난 12개월까지) | ~ | ~ | × | × | ○ | | | 37개국 중 23개국 |

주: 체크(✓)는 도구의 성능이 좋다는 증거, 물결(~)은 도구의 성능이 보통 수준이라는 증거, 엑스(X)는 도구의 성능이 좋지 않다는 증거, 동그라미(○)는 증거가 제한적이거나 누락되었다는 것을 나타낸다. 빈칸의 경우 해당 도구/주제에 대한 연구가 본 간행물에서 검토되지 않았음을 의미한다. * 독일이 1998년 독일 국가건강면접조사(German National Health Interview and Examination Survey)에 SF-36(단축형 척도인 SF-12가 아닌)을 포함시켰으나 라이선스 비용 문제로 인해 향후 해당 척도는 독일에서 사용되지 않는다. 각 도구에 대한 자세한 내용은 부속서 2.A 및 부속서 2.B를 참조한다. 국가 범위는 설문조사에 참여하지 않은 에스토니아를 제외한 모든 OECD 국가를 의미한다. 출처: 2022년 1월 각국 통계청에 발송된 OECD 설문지에 대한 응답에 기반하여 작성됨.

통계 품질

인구 정신건강에 적합한 측정 도구는 신뢰도, 타당도, 다양한 잠재변수(latent constructs)를 구별하는 능력, 무응답 또는 응답 거부 최소화, 집단 간 비교 가능성, 변화에 대한 민감도 등 다양한 통계 품질 특성에서 우수한 성능을 보여야 한다. 또한 응답자의 피로를 방지하기 위해 문항 간의 중복이 적고 설문 길이를 짧게 유지하는 등의 실질적인 사항도 고려하는 것이 중요하다. 이러한 다양한 통계 품질 특성은 상호 작용하므로 실제로 각각의 장단점을 균형 있게 고려하여 합리적인 해결책을 찾는 것을 목표로 한다. 어떤 도구가 한 가지 품질 기준(타당도)에서 우수한 성능을 보이더라도 다른 품질 기준(설문지 길이 또는 무응답률)에서는 저조한 성능을 보일 수 있다. 따라서 설문조사 제작 시 측정 항목을 선택하기 전에 각 접근 방식의 비용과 이점을 고려하여 상황에 적합한 도구를 파악하는 것이 중요하다.

정신건강에 대한 설문조사 측정을 어느 정도 신뢰할 수 있는가?

인구 정신건강 측정은 특정 상황에서 개인을 면담하거나 평가할 때 일관된 결과가 도출되어야 한다. 신뢰도라고 부르는 이러한 개념은 결과에서 감지된 변화가 도구 자체의 문제, 즉 측정 오류로 인한 문제일 가능성이 작고 대신 개인의 정신건강에 대한 실제 근본적인 변화를 반영하는지 확인하는 것이다(상자 3.1).

상자 3.1. 통계적 정의: 신뢰도

신뢰도의 두 가지 중요한 측면은 검사-재검사 신뢰도와 내적 일관성 신뢰도이다(OECD, 2013^[1], OECD, 2017^[2]).

검사-재검사 신뢰도는 시간이 경과해도 변하지 않는 척도의 안정성과 관련이 있다. 특정 설문 항목에 대해 일정 시간의 간격을 두고 재면담하거나 재평가하여 응답을 서로 비교한다. (측정 대상의 근본적인 상태에 변화가 없다고 가정할 때) 두 시기의 응답이 서로 높은 상관관계를 갖는다면 신뢰할 수 있는 측정치로 기대한다. 최초 면담과 후속 면담 사이의 시간 간격에 대한 정해진 규칙은 없으며 평가 유형에 따라 짧게는 2일~14일, 길게는 6개월까지의 범위에서 시행된다(NHS Health Scotland, 2008^[3]).

특정 개인의 정신건강 상태(특히 정서적 상태)는 단기간에 변동될 수 있기 때문에 정신건강 측정 도구의 경우 검사-재검사 기준을 신중하게 적용해야 한다. 즉 특정 증상이나 상태를 측정하는 도구는 신뢰도가 높을 수 있지만 증상과 경험 자체가 수일 또는 수주에 걸쳐 호전되거나 악화될 수 있기 때문에 동일한 개인에 대해 서로 다른 결과가 도출될 수 있다. 이에 따라 인구 정신건강 결과의 측정 측면에서 검사-재검사 신뢰도는 특히 다음과 관련이 있다.

- 개인이 정신건강 상태를 진단받은 적이 있는지 여부와 관련된 간단한 측정(우수한 도구는 검사-재검사 간의 높은 상관관계를 가짐)
- 동일한 응답자에게 시행 시 단축형 측정(또는 임상 진단에 대해 검증 중인 측정)과 확장형 측정(또는 임상 진단)이 검사-재검사에서 동일한 정확도로 수행되는지 여부를 확인
- 단기간, 대규모 표본을 대상으로 증상 기반 측정 척도의 광범위한 안정성을 확립하는 것, 즉 일련의 증상에 대한 문항의 검사-재검사 간의 상관관계가 특정 개인에게 안정적인 가능성은 낮은(증상 자체가 항상 안정적이지 않은 경우) 반면 개인 수준에서 증상의 일상적인 변동은 대규모 표본에서 사라져 단기간에 걸쳐 유사한 점수 분포가 나타날 것으로 예상된다².

따라서 검사-재검사 신뢰도 평가는 충분히 안정적이면서도 시간이 경과함에 따라 변화에 민감한 측정치 간의 균형을 의미한다. 검사-재검사 신뢰도에서 우수한 성과를 보이는 도구가 변화에 대한 민감도를 측정하는 검사에서는 성능이 저조할 수 있기 때문에 어떤 도구를 사용할지 결정할 때 통계 품질 특성을 전체적으로 살펴보는 것이 중요하다.

내적 일관성 신뢰도는 동일한 구성개념을 포착하는 것을 목표로 하는 경우 설문조사 도구 내의 개별 항목 간의 상관관계 정도를 평가한다. 인구 수준의 정신건강을 측정하는 맥락에서 내적 일관성 신뢰도는 우울증과 불안을 측정하도록 설계된 항목 모음에서 우울증 항목 간에 서로 상관관계가 있고 불안 항목 간에 서로 상관관계가 있음을 의미할 수 있다(요인적 타당도에 대한 논의는 상자 3.3 참조). 내적 일관성 신뢰도에 가장 널리 사용되는 계수는 총 문항 수, 개별 항목 쌍 간의 공분산, 전체 점수 분산의 함수인 크론바흐 알파(Cronbach's alpha)이다¹. 보편적인 합의는 없지만 이상적인 계수 값은 0.7~0.9 사이로 간주된다(NHS Health Scotland, 2008^[3]). 0.7 미만의 값은 척도 내 항목이 동일한 개념을 측정하지 못한다는 사실을 반영할 수 있으며 (OECD, 2013^[11]), 0.9 초과 값은 중복 항목이 있음을 의미할 수 있다.

주:

1. 크론바흐 알파 계수는 다중 항목 도구의 내적 일관성 신뢰도를 평가하기 위해 일반적으로 사용된다. 이 계수는 쌍을 이룬 항목의 평균 공분산에 척도에 포함된 총 항목 수를 곱하고 이 결과를 분산-공분산 행렬의 모든 요소의 합으로 나누어 계산한다(OECD, 2013^[11]). 그 결과 0(척도 항목이 서로 완전히 독립적이며 공분산 없음)에서 1(척도 항목이 겹치며 완전한 공분산)까지의 계수가 산출된다.
2. ‘단기간’의 정의는 주관적이며 상황에 따라 달라질 수 있다. 예를 들어 며칠의 기간이 검사-재검사 평가를 실시할 수 있는 허용 가능한 단기간으로 간주될 수 있으나 그 사이에 긍정적인든 부정적인든 극심한 충격이 발생할 경우 기본 분포가 변화할 수 있는 충분한 근거가 된다. 코로나19 팬데믹 기간에 정신건강에 대한 데이터를 높은 빈도로 수집한 결과, 봉쇄 정책의 도입 및 완화와 함께 많은 정서 기반 측정치가 큰 폭으로 급증하는 등 변동성이 나타났다.

신뢰도 측정에 대한 선별 도구의 성능은 도구의 유형과 측정 결과에 따라 다르며 심리적 고통의 일반적인 측정에 관한 연구에 따르면 다양한 결론이 도출되었다. 일반 건강 설문지(GHQ-12)와 단축형 건강 설문조사-36(SF-36) 및 그보다 하위 구성요소인 정신건강검사(MHI-5)는 신뢰도가 높은 것으로 나타났다 (Schmitz, Kruse 및 Tress, 2001^[4], Ohno 외, 2017^[5], Elovainio 외, 2020^[6], Strand 외, 2003^[7]), 그러나 케슬러 확장형(K10)이 내적 일관성 신뢰도가 나타난 것과 달리, 케슬러 단축형(K6)의 검사-재검사 신뢰도는 어떤 연구에서도 평가되지 않았다(EI-Den 외, 2018^[8], Easton 외, 2017^[9]).

반대로 특정 정신 상태, 특히 우울증 선별 도구에 대한 연구가 활발히 이루어지고 있으며 검사-재검사 신뢰도와 내적 일관성 신뢰도 측면에서 모두 신뢰할 수 있는 것으로 나타났다. 55개의 다양한 우울증 선별 도구에 대한 메타 분석 결과에 따르면 환자 건강 설문지(PHQ-9)가 가장 많은 평가를 받은 도구로 나타났으며, 많은 연구에서 PHQ-9와 PHQ-8(자살 충동에 관한 마지막 문항이 삭제된 단축형 버전) 모두 신뢰도와 타당도가 높은 것으로 나타났다(EI-Den 외, 2018^[8]). 그러나 동 연구에서 환자 건강 설문지-2(PHQ-2)의 타당도와 신뢰도에 대한 일관된 데이터가 부족하다는 점을 발견했다. PHQ-2를 평가한 6개의 보고서에서 내부 일관성 또는 검사-재검사 신뢰도에 대해 보고한 보고서는 단 한 건뿐이었다(EI-Den 외, 2018^[8]). 이로 인해 저자는 이용 가능한 데이터로 PHQ-2의 신뢰도를 확인할 수 없다고 경고했다. PHQ보다 연구가 부족 하긴 하지만, 역학연구 우울척도(CES-D)에 대한 연구에 따르면 두 측정 항목에서 우수한 신뢰도를 가지는 것으로 나타났다(Ohno 외, 2017^[5]). 불안 도구 중 범불안장애 선별 도구(GAD-7 및 GAD-2)는 검사-재검사 신뢰도 및 내적 일관성 신뢰도가 우수한 것으로 나타났다(Ahn, Kim 및 Choi, 2019^[10], Spitzer 외, 2006^[11]).

우울증과 불안의 복합 측정치를 산출하기 위해 PHQ-2와 GAD-2를 결합한 환자 건강 설문지-4(PHQ-4)에 대한 연구에 따르면 신뢰도가 낮긴 하지만, 여전히 수용 가능한 수준인 것으로 나타났다(두 하위 척도 모두에 대한 크론바흐 알파가 0.8 이상으로)(Kroenke 외, 2009_[12]). PHQ-4에 대한 다른 연구에서는 항목 간 상관관계가 낮은 것으로 나타났지만 척도의 길이가 짧다는 점을 고려할 때 신뢰도가 허용 가능한 수준인 것으로 판단했다(Löwe 외, 2010_[13])³. 크론바흐 알파는 부분적으로 전체 항목 길이의 함수이므로(상자 3.1 참조) 척도의 길이가 짧을수록 구성 개념 별 내적 일관성에 대한 검증 결과가 낮게 나타난다. 그러나 척도의 길이가 짧으면 문항 간 중복성이 적어져 응답자의 응답 부담이 적기 때문에 설문조사 작성자가 선호하는 경우가 많다.

긍정적 정신건강의 측면을 포착하는 복합 척도 역시 신뢰할 수 있는 것으로 밝혀졌다. 14개의 문항으로 구성된 워릭-에딘버러 웰빙지수(WEMWBS) 도구에 대한 연구에 따르면 해당 도구는 검사-재검사 신뢰도가 높고(일주일에 0.83), 크론바흐 알파가 높게 나타났다(약 0.9)(Tennant 외, 2007_[14], NHS Health Scotland, 2016_[15]). 그러나 저자들은 크론바흐 알파가 높은 것은 척도 항목에 일부 중복이 있음을 시사하며 이로 인해 7개 항목으로 구성된 단축형 버전(SWEMWBS)이 개발되었음을 강조했다(Tennant 외, 2007_[14], NHS Health Scotland, 2016_[15]). WEMWBS와 SWEMWBS에 대한 여러 연구에서 두 도구 모두 검사-재검사 신뢰도가 높은 것으로 나타났다(Stewart-Brown, 2021_[16], Shah 외, 2021_[17]). 세계보건기구-5(WHO-5) 복합 척도 역시 다양한 환경에서 신뢰도 테스트를 거쳤다(Dadfar 외, 2018_[18], Garland 외, 2018_[19]). 마찬가지로 MHC-SF의 검사-재검사 신뢰도는 중간 정도에 불과하지만 내적 신뢰도가 높은 것으로 나타났다(Lamers 외, 2011_[20]).

정신건강 상태에 대한 일반적인 자기보고 지표의 신뢰도를 조사한 연구는 거의 없지만 미국의 한 연구에 따르면 해당 척도가 수용 가능한 검사-재검사 신뢰도를 가지고 있음을 시사한다. 미국 질병통제센터(Centers for Disease Control)에서 사용하는 건강 관련 삶의 질 측정 도구인 행동위험요인감시시스템(BRFSS) 설문 조사는 신체적 건강과 정신적 건강을 결합하여 지각된 건강을 측정한다. 미국 미주리주(Missouri)에서 실시된 연구에 따르면 4개 항목으로 구성된 단축형 버전의 도구는 수용 가능한 수준의 검사-재검사 신뢰도를 가지고 있으며 내적 타당도가 높으나 노년층에서 신뢰도가 낮은 것으로 나타났다(Moriarty, Zack 및 Kobau, 2003_[21]).

상자 3.2. 주요 메세지: 신뢰도

- 특정 정신장애를 식별하는 설문조사와 긍정적 정신건강을 식별하는 설문조사를 포함한 대부분의 정신 건강 선별 도구는 검사-재검사 신뢰도 및 내적 일관성 신뢰도를 측정할 결과 신뢰도가 높은 것으로 나타났다.
- 검사-재검사 신뢰도는 품질 기준으로 맹목적으로 적용하기보다는 시간이 경과하면서 나타나는 변화에 대한 측정 민감도와 함께 고려해야 한다.
- 선별 도구(특히 우울증에 초점을 맞춘 도구)의 신뢰도를 뒷받침하는 강력한 근거가 있으며 이보다는 미약하지만 긍정적 정신건강 복합 척도의 신뢰도에 대한 근거도 있다. 그러나 정신건강 상태에 대한 일반적인 자기보고 지표의 신뢰도를 평가한 연구가 충분히 이루어지지 않았으므로 앞으로 해당 분야에 대한 활발한 연구가 필요할 것이다.

도구가 목표한 결과를 정확하게 측정하는가?

우수한 측정 도구는 신뢰도 외에도 타당도가 있어야 한다. 즉 도구를 통해 도출된 측정치는 기본 개념을 정확하게 반영해야 한다. 더 객관적인 지표의 경우, 동일한 구성 개념에 대해 객관적으로 측정된 자기보고식 측정치와 비교해 타당도를 평가할 수 있다. 예를 들어 이론적으로 응답자의 자기보고 소득은 세금 신고서 또는 급여 명세서와 교차 확인하여 응답이 정확하게 보고되었는지 확인할 수 있다. 물론 이를 체계적으로 수행할 수 없는 현실적인 이유가 존재하지만, 이는 자기보고 소득 데이터의 타당도를 평가하는 방법을

보여준다. 이와 대조적으로 주관적 웰빙, 신뢰 또는 실제 정신건강과 같은 주관적 지표의 ‘객관적 사실’을 확인하는 것은 불가능하다. 이는 타당도를 평가할 수 없다는 의미는 아니다. OECD 측정 지침은 주관적 지표의 타당도를 평가하기 위해 안면 타당도, 수렴 타당도, 구성 타당도 개념을 사용한다(OECD, 2013^[1], OECD, 2017^[2])(상자 3.3).

많은 주관적 지표와 달리 정신건강을 평가하는 대부분의 선별 도구는 일반적인 정신장애의 진단을 위한 면담과 비교해 타당화되어 정확성과 실제 의미를 엄격하게 평가할 수 있다. 정신건강 선별 도구의 타당도를 검증하는 가장 일반적인 진단 면담은 세계보건기구의 복합 국제진단 면담 도구(WHO-CIDI)로 임상 및 연구 목적뿐만 아니라 역학 연구용으로도 사용하도록 설계되었다(자세한 내용은 제2장 참조). 해당 도구를 사용하여 정신장애의 유병률, 장애의 중증도, 가정 관리, 일과 삶의 균형, 인간 관계 및 사회 생활에 미치는 영향, 정신건강 서비스 이용 및 약물 사용 등을 측정할 수 있다. CIDI는 정신건강 설문조사의 항목을 평가하는 표준 기준으로 널리 인식되고 있으나 이에 대한 비판과 타당도 문제가 존재한다(상자 3.4).

상자 3.3. 통계적 정의: 타당도

타당도 신뢰도보다 검증하기 어려우며, 특히 객관적 사실을 알 수 없고 일반적으로 객관적 측정치와 동등하게 비교할 수 없는 주관적인 자료에 대한 타당도 검증은 특히 어렵다. 주관적 측정치의 타당도를 평가하는 방법에는 안면 타당도, 수렴 타당도 및 구성 타당도 등 세 가지 방법이 있다.

안면 타당도는 지표가 응답자와 (잠재적) 데이터 사용자에게 직관적으로 이해되는지 여부를 평가한다. 안면 타당도는 무응답을 통해 간접적으로 측정할 수 있다. 무응답률이 높으면 응답자가 문항의 관련성이나 유용성을 이해하지 못하거나 인식하지 못하고 있음을 나타낼 수 있다. 정신건강의 경우 높은 무응답률은 안면 타당도의 결여보다는 낙인과 편견으로 인해 해당 주제에 대해 응답을 꺼리는 심리가 반영되었을 수도 있다(정신건강 측정 및 무응답에 대한 자세한 논의는 본 장의 후반부에서 다룸). 인지 면담(Cognitive interviewing)도 사용될 수 있다.

수렴 타당도는 동일한 기저의 결과를 상이한 방법으로 측정했을 때 측정치의 상관관계에 따라 평가된다. 정신건강 도구를 예로 들어 연구자가 불안을 평가하는 새로운 도구를 도입하려는 경우 불안에 대한 데이터, 진단 또는 정신건강 관련 의료 서비스 이용, 불안 수준에 대한 자기보고 평가, 스트레스 및 불안의 생체-신체적 지표(심박수, 혈압, 신경영상 등)를 기존의 선별 도구와 비교하여 수렴 타당도를 검증할 수 있다.

구성 타당도는 지표가 기존의 이론이나 연구와 일치하는 정도를 의미한다. 예를 들어 연구에 따르면 정신 건강과 신체건강은 서로 상관관계가 있으며 함께 움직인다. 따라서 새로운 정신건강 도구가 신체건강과의 상관관계가 거의 없거나 해당 도구로 측정된 정신건강의 변화가 신체건강의 변화를 반영하지 않는다면 해당 척도는 구성 타당도가 낮다고 간주할 수 있다. 건강의 사회적 결정 요인에 관해 증가한 문헌도 이와 유사한 방식으로 구성 타당도를 평가하는데 활용할 수 있다.

위에서 언급한 타당도의 세 가지 측면 외에도 정신건강 설문조사 항목의 임상적 타당도 검사는 종종 준거 타당도, 요인 타당도 및 교차 타당도 등의 세 가지 평가를 추가로 참조하는 경우가 많다.

준거 타당도는 항목을 비교할 수 있는 표준 기준이 있을 경우에만 평가가 가능하다. 정신건강의 경우 이러한 표준은 일반적으로 구조화된 면담(예: CIDI, 부속서 2.B 참조)이다. 준거 타당도는 측정의 심리측정적 속성, 즉 측정이 표준 기준과 어떻게 비교되는지를 평가한다. ‘진양성(true positive)’(예: 설문조사에서 우울증의 위험군을 얼마나 자주 정확하게 식별하는지)을 정확하게 식별할 수 있으면 민감도가 높고 ‘진음성(true negative)’(예: 우울증 *비*위험군을 정확하게 식별하는지)을 정확하게 식별할 수 있으면 특이도가 높다고 간주할 수 있다.

진단 정확도를 설정하기 위해 민감도와 특이도를 다양한 임계값에서 ROC(receiver operating characteristic) 곡선으로 표시한다. 그런 다음 곡선 아래 면적(AUC)을 사용하여 표준 기준과 비교하여 선별 도구의 진단 성능을 평가할 수 있다¹.

요인적 타당도는 여러 항목으로 구성된 설문조사 도구가 하나의 기본 개념을 측정하는지 또는 여러 개념을 측정하는지 평가한다. 거의 모든 경우에서 단일한 개념만을 평가하는 경우 단일차원성을 특징으로 하며, 이는 정신건강 도구가 우울증, 불안 또는 잠재적 웰빙을 측정하고 있음을 보장한다. 그러나 척도가 정신건강의 여러 측면을 평가하는 경우에는 다차원성이 필요하다. 예를 들어 우울증과 불안을 측정하는 PHQ-4의 요인 평가는 실제로 두 가지 잠재요인을 식별한다(Löwe 외, 2010^[13]). 요인적 타당도는 일반적으로 확인적 요인분석(CFA) 또는 탐색적 요인분석(EFA)을 사용하여 평가한다. 전자의 경우 연구자는 관찰 변수(예: PHQ-8 도구에 대한 응답자의 답변)와 잠재변수(예: 우울증) 간의 관계가 주어진 모형에 적합하다는 가설을 검증한다. 즉 CFA를 사용하여 관찰된 데이터 세트에 특정한 수의 잠재요인이 있다는 가설을 검증한다. 연구자는 EFA를 사용하여 이론적 모형을 강조하는 대신 거꾸로 작업하여 기본 요인 구조를 도출한다(Suhr, 2006^[22]).

교차 타당도 또는 문화 타당도는 측정이 여러 인구집단에 걸쳐 적용되는 정도를 의미한다. 교차 타당도는 증상이 표현되는 방식에 영향을 미치는 문화적 요인, 임상적 편향(암묵적 또는 명시적), 응답자의 언어적 제한(모국어가 아닌 다른 언어로 된 도구를 시행하는 경우) 및 응답 행동의 차이(예: 리커트 척도에서 극단값보다 중간값을 선택할 가능성이 높음)을 포함하여 다양한 방식으로 정신건강 결과 측정을 편향시킬 수 있다(Leong, Priscilla Lui 및 Kalibatseva, 2019^[23]). 교차 타당도는 설문조사 도구를 맹목적으로 적용하는 것보다 필요한 인구집단에서 검증함으로써 확인된다.

주:

1. 수신자조작특성곡선(receiver operating characteristic curve, ROC)은 진음성률에 대해 진양성률을 표시하여 진단 능력을 시각화한다. 이 곡선은 유형 1(가양성) 및 유형 2(가음성) 오류를 모두 최소화하는 최적의 준거점수를 결정하는 데 사용할 수 있다. ROC 분석은 준거점수 기준값을 결정하는 데 사용되며, 본 장의 후반부에서 논의한다. 임상심리학에서의 ROC 사용에 대한 자세한 내용은 Pintea 및 Moldovan(2009^[24])과 Streiner 및 Cairney(2007^[25])를 참조한다.

선별 도구의 타당도를 평가하기 위해 연구자는 일반적으로 응답자가 자기보고 척도에 응답하고 구조화된 CIDI 면담에 참여하여 도출된 두 가지 응답을 비교하는 연구를 시행한다. 민감도와 특이도가 높은 선별 도구는 준거 타당도가 높은 것으로 간주된다. 준거 타당도는 선별 도구가 CIDI의 진단 결과를 반영하도록 설계되었음을 보장하지만 선별 도구는 설계상 특정 정신장애에 대해 높은 유병률을 추정한다(상자 2.1 참조). 수렴 타당도는 새로운 우울증 측정 도구가 기존의 우울증 측정 도구와 유사한 성능을 보이는지 확인하기 위해 다양한 선별 도구를 서로 비교하여 평가한다. 이 접근 방식은 단축형 선별 도구를 검증할 때 사용되는 경우가 많으며 단축형 설문조사가 확장형 설문조사만큼의 성능을 보이는지 여부를 확인하기 위해 사용된다. 본 장에서 설명된 대부분의 선별 도구는 일반적인 정신장애를 진단하기 위한 면담과 비교해 검증되었으며 연령대, 성별 및 사회경제적 지위에 걸쳐 우수한 심리측정적 속성(높은 민감도 및 특이도)이 보고되었다(Gill 외, 2007^[26], O'Connor 및 Parslow, 010^[27], Huang 외, 2006^[28])(상자 3.3).

상자 3.4. 구조화된 면담의 타당도

선별 도구를 검증하기 위해 구조화된 면담을 사용할 때 주의해야 할 한 가지 중요한 점은 구조화된 면담이 ‘진실된(true)’ 정신건강을 정확하게 측정할 수 있다고 전제한다는 점이다. 이는 다음의 두 가지 상황에서 문제가 제기된다. (1) OECD 국가에서 사용되는 대부분의 선별 도구는 최신 버전인 2013년에 발간된 정신질환 진단 및 통계 편람 제5판(DSM-5)이 아닌 1994년에 발간된 정신질환 진단 및 통계 편람(DSM-IV) 제4판(현재는 구 버전)을 기준으로 검증되었다는 점, (2) 정신질환 진단 및 통계 편람(DSM) 자체가 문화 간 정확한 진단 데이터를 제공하는 정도에 대한 의문이다.

구조화된 면담의 타당도와 관련된 첫 번째 우려사항은 일반적으로 사용되는 선별 도구가 최신 버전의 정신질환 진단 및 통계 편람(DSM)에 대해 검증되지 않았다는 사실과 관련 있다(Statistics Canada, 2021_[29]). 전체적으로 정신질환 진단 및 통계 편람 제4판(DSM-IV)과 제5판(DSM-5) 간에는 총 464개의 차이점이 존재한다. 대체로 정신질환 진단 및 통계 편람 제5판(DSM-5)에는 이전에 분리되어 있던 많은 장애가 여러 가지 특징이나 증상을 공유하기 때문에 더 적은 수의 진단 범주가 포함되어 있다. 또한 가능한 한 개인의 기능 상태와 진단을 분리하여 작성하였다. 앞으로 영향을 미칠 수 있는 한 가지 영역은 범불안장애의 진단 기준이 낮아진 것인데, 이전에 일상적인 걱정으로 여겨지던 것을 병리화한다는 비판을 받는 변화이다(Murphy 및 Hallahan, 2016_[30]).

요약하자면 정신질환 진단 및 통계 편람(DSM) 개정 시 진단 범주가 재구성되고 일부 진단 기준이 업데이트되는 등 변경사항이 있지만, 서로 다른 버전 간에는 설계상 연속성이 있으며 대부분의 변경사항은 경미한 수준이다. 이와 상관없이, 최근의 임상 관행을 최신 상태로 유지하기 위해서는 CIDI와 같은 도구의 업데이트를 통해 이를 달성할 수 있을 것이다.

두 번째 우려사항은 이러한 진단 검증을 미국 이외의 지역 및 인구집단에 적용할 수 있는지에 대한 것이며, 최소한 검증 연구가 서로 다른 지역적 맥락에서 수행되어야 한다. 이 외에도 검증을 뒷받침하는 임상 진단 자체에 결함이 있는 경우 지리적으로 다양한 임상 환경에서 정신건강 선별 도구를 검증하는 것은 불충분할 수 있다. Haroz와 동료들은 77개의 국적과 민족을 반영하는 우울증에 대한 질적 연구 138건을 검토하여 문화 간 편견 정도를 조사했다(Haroz 외, 2017_[31]). 그 결과 비서구권 인구에서 가장 자주 언급된 우울증의 특징 15개 중 7개만이 정신질환 진단 및 통계 편람 제5판(DSM-5)의 주요우울장애 진단을 반영한다는 사실을 발견했다. ‘집중력 문제’ 및 ‘정신운동 초조 또는 지연’ 등 DSM에 명시된 진단적 특징은 자주 나타나지 않았으며, 진단 기준에 포함되지 않는 ‘사회적 고립 또는 외로움’, ‘울음’, ‘분노’, ‘일반적 통증’을 포함하는 특징은 자주 나타났다. 일부 특징은 특정 지역에서 더 자주 관찰되었다. 남아시아와 동남아시아에서는 ‘걱정’이, 동남아시아와 사하라 이남 아프리카에서는 ‘생각이 너무 많음’이 자주 관찰되었다. 이는 PHQ-9 또는 GAD-7과 DSM 기준의 긴밀한 일치성으로 인해 더 길고 집중적인 선별도구 또는 구조화된 면담에 비해 내재된 목표 구성 개념(예: 우울증, 불안)을 발견하는데 제한이 있을 수 있음을 의미한다(Ali, Ryan 및 De Silva, 2016_[32], Sunderland 외, 2019_[33]).

DSM 기준에 대한 비판이 존재하지만, DSM은 여전히 정신건강 결과에 대해 국가 간 비교 가능한 데이터를 생산하는 데 가장 유용한 도구로 여겨진다. DSM 후속 개정판 초안 작성 시 문화 타당도에 대한 고려 사항을 포함하며 개선할 수 있다.

일부 OECD 국가에서는 정신건강에 대한 정의를 서구적 관점에 뿌리를 둔 전통적인 임상심리학적 관점을 넘어 더욱 광범위한 관점을 포괄하도록 확장했다. 예를 들어 뉴질랜드에서는 정신건강 및 중독에 대한 정부 조사에서 정신건강에 대한 마오리족 관점이 포함되어 있다(New Zealand Government, 2018_[34]). 이와

유사한 맥락에서 스웨덴 정부는 향후 정신건강에 대한 전략의 초안을 작성할 때 사미족 의회의 의견을 요청할 예정이다(Public Health Agency of Sweden, 2022^[35]).

일반적인 심리적 고통에 대한 측정 전반에서 케슬러 및 MHI-5 척도가 GHQ-12보다 준거 타당도가 높은 것으로 나타났다. 연구에 따르면 K10 및 K6 척도는 GHQ-12보다 우울 및 불안장애를 감지하는 데 있어 심리측정적 속성(신뢰도와 타당도 모두 포함)과 전반적인 변별력이 더 우수한 것으로 나타났다(Furukawa 외, 2003^[36], Cornelius 외, 2013^[37]). 또한 SF-12 도구의 정신건강 구성요소는 GHQ-12에 비해 특정 정신건강 상태의 유무를 구별하는 성능이 우수하다(Gill 외, 2007^[26]). MHI-5 도구는 주요 우울증 및 불안장애를 포함한 여러 정신건강 상태를 평가하는 데 MHI-18 및 GHQ-30만큼 유효한 것으로 나타났으나 전체 정동장애에 대해서는 MHI-18보다 성능이 떨어지는 것으로 나타났다(Berwick 외, 1991^[38]). MHI-5는 일반적인 도구로 설계되었지만, 우울증 과/또는 불안에 대한 특정 위험을 식별하는 데 효과적인 것으로 입증되었다(Yamazaki, Fukuhara 및 Green, 2005^[39], Rivera-Riquelme, Piqueras 및 Cuijpers, 2019^[40]).

우울증의 중증도를 진단하고 등급을 매기는 데 사용되는 도구의 민감도와 특이도에 대한 최근 메타 분석에 따르면 평균적으로 PHQ-9가 CES-D를 포함한 여타 선별 도구에 비해 가장 높은 민감도와 특이도를 보인 것으로 나타났다(Pettersson 외, 2015^[41]). CDC 행동위험요인감시시스템(BRFSS) 설문조사에서는 다른 버전의 PHQ-8이 사용되었다. 이 척도는 PHQ-8 일수(PHQ-8 days)라고 불리며, 응답자에게 지난 4주 동안 PHQ-8을 구성하는 8가지 우울 증상을 각각 경험한 일수에 대해 작성하도록 요청한다. 이를 통해 0점~112점 범위의 척도를 산출하며 경미한 우울증에 걸릴 위험이 있지만 현재 정신적 웰빙 수준이 높은 개인의 우울증 위험을 보다 세밀하게 살펴볼 수 있으며 변화에 대한 민감도가 높은 것으로 나타났다(Dhingra 외, 2011^[42]). PHQ-2는 내적 일관성, 구성 타당도 및 상관관계 수렴 타당도에 대해 평가되었으나 메타 분석에서 준거 타당도에 대한 평가 근거는 발견되지 않았다(El-Den 외, 2018^[8]). 또 다른 개요에서는 PHQ-2가 노인, 임산부 또는 산후 여성, 관상동맥심장질환 또는 에이즈(HIV/AIDS)와 같은 특정 질환을 앓고 있는 환자 등 특정 인구집단에 대해 준거 타당도가 우수하다는 근거를 인용했다(Löwe 외, 2010^[13]).

우울 증상에 대한 자기보고 척도는 타당도가 높은 편이지만 불안장애에 대한 척도는 임상 인구에서 민감도 및 특이도가 다소 저하된 것으로 나타났다. 연구에 따르면 이는 다양한 유형의 불안장애가 우울장애보다 이질적인 증상을 더 많이 가지고 있기 때문일 수 있다(Rose 및 Devine, 2014^[43]). 그럼에도 불구하고 GAD-7과 GAD-2는 여러 연구에서 검증되었다. GAD-7은 범불안장애에 대한 간략한 임상적 척도를 제공하기 위해 설계되었으며 타당도 검증 결과 타당도(준거, 구성, 요인 등)가 우수한 것으로 나타났다. 또한 GAD-7과 PHQ-8의 요인 타당도 평가에 따르면 불안 척도와 우울증 척도 간의 상관관계가 높음(0.75)에도 불구하고 두 척도가 중복되지 않고 상호 보완적인 것으로 나타났다. 불안 수준이 높은 환자의 절반 이상에서 우울증 수준이 높지는 않았다(Spitzer 외, 2006^[11]). GAD-7과 다른 두 가지 불안 척도와 높은 상관관계는 수렴 타당도가 우수함을 나타낸다(Kroenke 외, 2007^[44], Spitzer 외, 2006^[11])⁴. GAD-7과 GAD-2는 불안장애의 네 가지 주요 형태인 범불안장애, 공황장애, 사회불안장애 및 외상 후 스트레스 장애를 감지하는 성능이 우수한 것으로 나타났다(Kroenke 외, 2007^[44]).

PHQ-4는 우울증과 불안의 위험성을 함께 측정하는 유효한 도구인 것으로 나타났다. 위에서 언급한 바와 같이 PHQ-4의 구성요소인 PHQ-2 및 GAD-2는 진단기준 표준면담에 대해 검증되었다(상기 언급된 바와 같이 PHQ-2 준거 타당도의 광범위한 적용 가능성에 대한 주의사항 포함). 연구에 따르면 PHQ-4 점수는 SF-20 기능적 상태 척도 및 사용한 장애 일수 등의 건강 정보와 연관되어 수렴 타당도 및 구성 타당도에

대한 근거를 제공한다. 또한 요인 분석에 따르면 PHQ-4는 두 개의 개별 요인이 있는 2요인 구조를 가지고 있으며, 이는 우울증과 불안장애를 모두 포착하는 것으로 나타났다(Löwe 외, 2010_[13]).

긍정적 정신건강의 측면을 포착하는 복합 척도 역시 타당도가 높은 것으로 나타났다. WEMWBS는 긍정적 정서를 포착하는 여타 척도와 높은 상관관계를 보이며 준거 타당도 및 수렴 타당도가 우수한 것으로 나타났다. WEMWBS와 WHO-5는 높은 상관관계를 보였으며(상관계수 0.77)(NHS Health Scotland, 2016_[15]) 신체적 건강이나 심리적 고통에 중점을 둔 기타 정신건강 척도(GHQ-12 포함)와는 상관관계가 다소 낮았다. WEMWBS에 대한 또 다른 연구에서는 단축형 버전이 확장형 버전과 높은 상관관계를 보여 14개 문항으로 구성된 확장형 버전보다 효율적이고 신속한 대안이 될 수 있다는 사실을 발견했다(Stewart-Brown 외, 2009_[45]). 라쉬 분석(Rasch analysis)에 따르면 WEMWBS는 문항 길이에도 불구하고 하나의 기본 요소를 가진 단일차원성 구성으로 나타났다(Stewart-Brown, 2021_[16])⁵. 여러 연구에 따르면 MHC-SF는 수렴 타당도가 우수하지만(Guo 외, 2015_[46], Petrillo 외, 2015_[47], Lamers 외, 2011_[20]) 덴마크의 인지 면담에서는 특히 사회적 하위 척도 문항에 대해 안면 타당도가 낮은 것으로 나타났다(Santini 외, 2020_[48]).

WEMWBS와 WHO-5는 긍정적 정신건강을 측정하기 위해 설계되었으나 우울증 과/또는 불안을 효과적으로 선별할 수 있는 것으로 나타났다. 한 연구에 따르면 WHO-5는 임상 환경에서 우울증 환자를 식별하는 데 민감도는 높지만 특이도는 낮은 것으로 나타났다(Topp 외, 2015_[49]). SWEMWBS에 대한 한 연구에 따르면 SWEMWBS는 PHQ-9($\rho = 0.6-0.8$) 및 GAD-7($\rho = 0.6-0.7$)과 상대적으로 높은 상관관계를 보여 일반적인 정신장애(CMD)를 측정하는 데 적합한 도구임을 시사한다. 그러나 정신건강이 낮은 수준의 개인을 식별하고 구별하는 데 있어 다른 도구의 민감도가 더 높을 수 있다(Shah 외, 2021_[17]). 다차원 문항반응이론을 통해 WEMWBS와 GHQ-12를 비교한 연구에 따르면 두 도구 모두 동일한 구성개념을 측정하는 것으로 나타났다(Böhnke 및 Croudace, 2016_[50]).

자기보고 정신건강(SRMH) 지표는 여타의 정신건강 척도와는 별개이기는 하지만 정신건강에 대한 타당화된 임상 척도와 비교한 결과 관련성이 있는 것으로 나타났다. SRMH는 케슬러 척도, PHQ 및 SF-12의 정신건강 구성요소와 상관관계가 있으며, 기타 정신건강 선별 도구의 검증 과정에서 수렴 타당도를 검증하는 데 사용되는 경우가 많다. 또한 SRMH는 신체건강 악화 및 의료 서비스 이용 증가와 관련이 있지만, 연구에 따르면 SRMH와 선별 도구 사이의 상관관계는 중간 정도이며, 이는 두 도구가 다소 상이한 현상을 포착하고 있음을 시사한다(Ahmad 외, 2014_[51]). 이는 향후 추가 연구가 필요하지만 자기보고 신체건강에 대한 종단연구 결과에서 일부 조명하고 있다. SRMH 척도는 다른 지표에 비해 사망률, 동반이환률 및 의료 서비스 이용에 대한 강력한 예측변수로 나타났으며, SRMH는 선별 도구에서 아직 나타나지 않는 정신건강 문제를 포착할 수 있는 것으로 나타났다(Ahmad 외, 2014_[51]). 반대로 신체건강과 정신건강을 모두 측정하는 건강관련 삶의 질(HRQoL)은 SF-36 척도와 구성 타당도 및 준거 타당도가 유사한 수준으로 나타났다(Moriarty, Zack 및 Kobau, 2003_[21]).

정신건강 선별 도구에 대한 연구에서는 단일 항목으로 구성된 정신건강 문항이 충분히 타당한지에 대해 상반되는 근거를 제시했다. MHI-5와 MHI-18의 성능을 비교 평가한 연구(단축형 버전에 유리한 결론이 도출됨)에 따르면 “귀하는 얼마나 자주 낙담하거나 우울한 기분이 들었습니까?”라는 단일 문항만으로도 주요 우울증을 감지하는 데 있어 MHI-5, MHI-18 및 GHQ-30과 마찬가지로 우수한 성능을 보이는 것으로 나타났다(Berwick 외, 1991_[38]). 그러나 초단기 선별 도구를 평가한 연구에 따르면 두 가지 문항으로 구성된 경우, 단일 문항보다 우울증을 선별하는 데 더 우수한 성능을 보이는 것으로 나타났다(Löwe 외, 2010_[13]). 반대로 코로나19 팬데믹 기간에 시행된 호주 국민정신건강측정(Australian Taking the Pulse of the Nation, TPPN) 설문조사에 따르면 K6에 비해 단일 항목으로 구성된 정신건강 척도의 심리측정적 속성이 우수한 것으로 나타났다. 단일 항목 측정은 항목 간 상관관계가 높고($\rho = 0.82$), 심리적 고통에 대한 민감도가 높은 것으로 나타났다(Botha, Butterworth 및 Wilkins, 2021_[52]).

상자 3.5. 주요 메세지: 타당도

- OECD 회원국에서 일반적으로 사용하는 모든 선별 도구는 임상 환경에서 타당화 되었으며, 수렴 타당도, 구성 타당도 및 준거 타당도가 높은 것으로 나타났다.
- 또한 긍정적 정신건강의 측면을 포착하는 복합 척도는 심리측정적 속성이 우수한 것으로 나타났으며, 우울증 또는 불안과 같은 특정 정신건강 상태를 선별하는 데 효과적 도구임이 입증되었다.
- 준거 타당도는 설문조사 도구 성능에 따라 임상 진단 면담의 표준 기준과 비교하여 평가되지만, 이는 임상 진단의 타당도를 전제로 하며, 모든 상황에 적용되지 않을 수 있다.

무응답률이 낙인에 대해 시사하는 바는 무엇이며, 정신건강 데이터의 집단 간 비교 가능성에 어떤 영향을 미치는가?

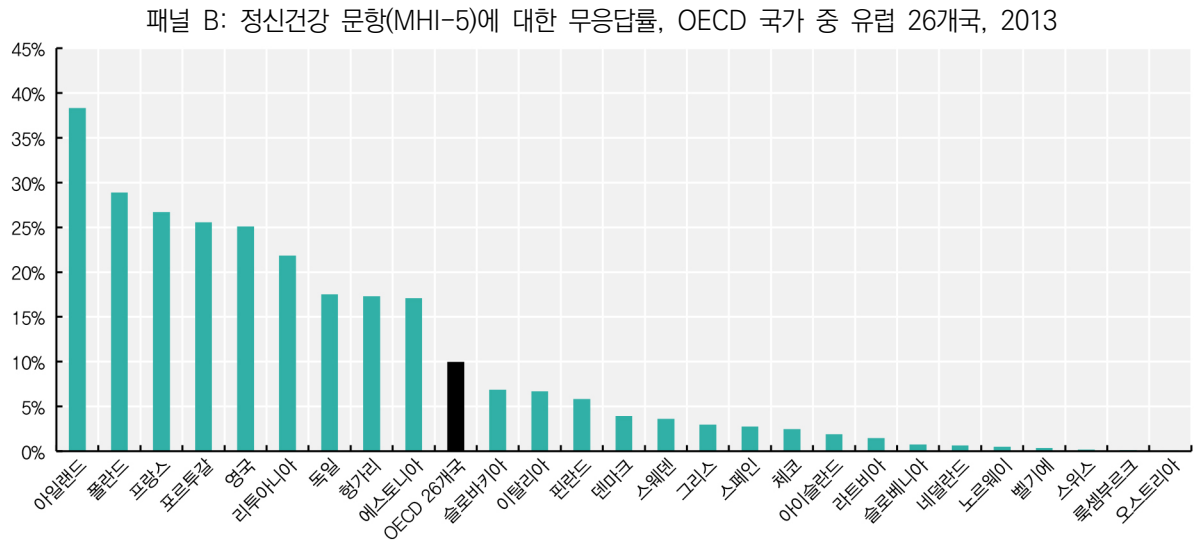
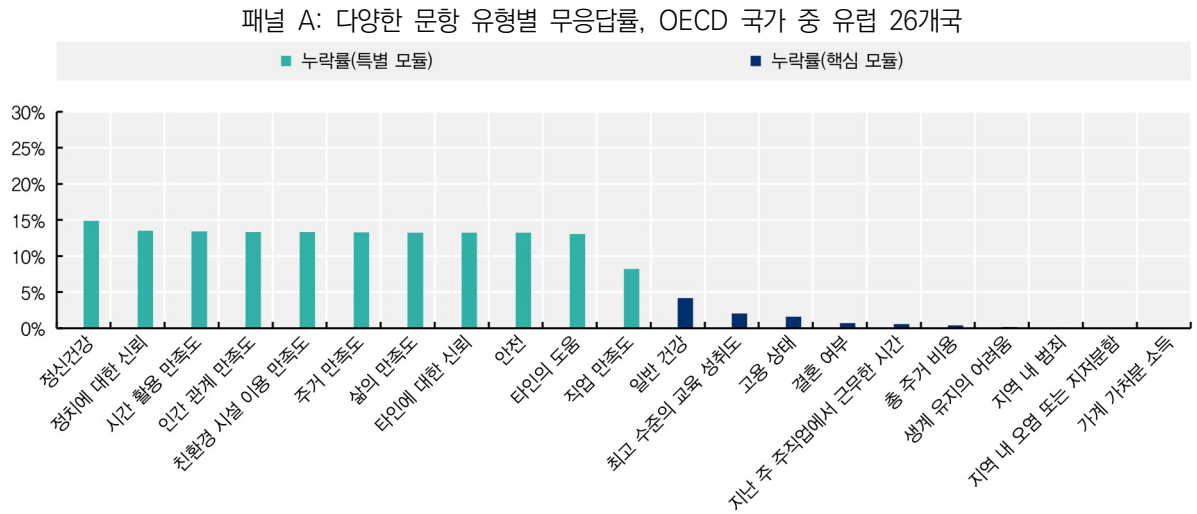
정신질환과 관련된 낙인은 정신건강 상태를 사실과 다르거나 축소하여 보고하는 것으로 이어질 수 있다(Hinshaw 및 Stier, 2008^[53]). 또한 정신건강에 대한 이해능력이 낮으면 개인이 경험한 증상을 기저 질환의 대표적 증상으로 인식하지 못하여 과소 보고로 이어질 수 있다(Tambling, D’Aniello 및 Russell, 2021^[54], Dunn 외, 2009^[55], Coles 및 Coleman, 2010^[56])⁶. 정신건강 상태에 대한 낙인은 모든 OECD 국가에서 여전히 중요하지만 국가 간 차이가 큰 편이다. 2019년에 실시된 설문조사에 따르면 OECD 19개국에서 응답자의 40%가 정신질환이 다른 질병과 동일하다고 생각하는지 여부를 묻는 문항에 동의하지 *않는다*고 답했으며 정신장애 병력이 있는 사람은 공직선거에 입후보를 금지해야 한다는 데 동의한 응답자가 4분의 1에 달했다(OECD, 2021^[57]). 낙인으로 인해 응답자는 정신건강 설문조사에 응답 시 실제 상태를 숨기거나 애초에 참여하지 않을 수 있다. 민감한 주제에 대한 설문조사를 실시할 시 데이터 기밀성을 명확히 보장하고 가족 구성원이 들을 수 없는 사적인 장소에서 면담을 진행하면 응답 거부 가능성을 최소화할 수 있다(Singer, Von Thurn 및 Miller, 1995^[58], Krumpal, 2013^[59])⁷.

심리적 고통이나 특정 정신장애를 경험하는 사람들은 설문조사 참여를 거부할 가능성이 더 높다는 근거가 있으며 이러한 무응답 편향으로 인해 정신질환의 전반적인 유병률을 과소평가하게 된다(de Graaf 외, 2000^[60], Eaton 외, 1992^[61], Mostafa 외, 2021^[62]). 최근 한 연구에서는 심리적 고통이 여러 경제적 전환(예: 실직)에 미치는 영향에 대해 GHQ-12의 기존 응답을 기반으로 산출한 점수와 잘못 보고된 응답을 조정하여 산출한 점수를 비교했다. 결과에 따르면 기존 응답을 기반으로 산출한 점수에서는 심리적 고통이 급여가 높은 일자리로의 전환과 높은 교육 성취도에 미치는 영향을 과소평가한 것으로 나타났다(Brown 외, 2018^[63]). 따라서 심리적 고통의 증상을 잘못 보고하면 편향되고 일관되지 않은 추정치가 도출될 수 있다. 그러나 모든 연구에서 동일한 결론에 도달한 것은 아니다. 미국공존질환조사(National Comorbidity Survey Replication, NCS-R)에 대한 연구에서는 무응답이 장애 유병률의 과소평가로 이어진다는 증거를 발견하지 못했다(Kessler 외, 2004^[64]).

2013년 유럽연합 소득 및 생활조건 통계(EU-SILC) 조사에 따르면 정신건강 관련 문항에 대한 무응답률은 15%로 높지만 다른 주관적 문항(예: 정치에 대한 신뢰도 13%, 직업 만족도 8% 등)에 대한 무응답률과 여전히 비슷한 수준으로 나타났다(그림 3.1 패널 A). 정신건강에 대한 무응답률(MHI-5를 통해 측정됨)이 높은 것은 부분적으로 유럽연합 소득 및 생활조건 통계(EU-SILC) 조사가 시행되는 방식에서 비롯된 것일 수 있다. 매년 핵심 모듈(2013년에는 웰빙에 초점을 맞춘 모듈) 외에 특정 주제에 대한 추가 문항이 포함된 특별 모듈이 시행되는데, 이는 일부 응답자가 이전 조사에 포함되지 않은 문항에 응답하는 데 문제가 있을 수 있음을 의미한다.

정신건강 문항에 대한 누락률과 무응답률은 국가마다 큰 차이를 보인다(그림 3.1, 패널 B). 2013년 유럽 연합 소득 및 생활조건 통계(EU-SILC) 조사에서 정신건강 모듈의 누락률은 아일랜드, 폴란드, 프랑스, 포르투갈, 영국, 리투아니아에서 20%를 넘었으나 노르웨이, 벨기에, 스위스, 룩셈부르크, 오스트리아에서는 1% 미만에 불과하였다.

그림 3.1. 정신건강 관련 문항의 경우 다른 문항에 비해 무응답률이 높으며, 이는 국가별로 큰 차이를 보인다



주: 이 수치는 설문조사 참여에 동의한 응답자 중 개별 문항에 응답하지 않기로 선택한 응답자만 포함되며, 전체 설문조사 참여를 거부한 사람은 고려되지 않는다. 응답자가 MHI-5의 개별 항목 5개 중 4개 이상을 거부하거나 '모르겠다'고 답한 경우, 정신건강 데이터가 누락된 것으로 간주된다. 특정 도구에 대한 자세한 내용은 부속서 2.B를 참조한다.

출처: 유럽연합 소득 및 생활조건 통계(EU-SILC) (n.d.^[65]) (데이터베이스), <https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions>.를 기반으로 OECD에서 계산함.

StatLink <https://stat.link/s3v124>

표 3.2는 국가별 무응답률의 차이가 낙인 수준과 관련이 있음을 시사하는 몇 가지 근거를 보여준다. OECD 국가 중 유럽 9개국에서 우울장애 유병률(PHQ-8로 측정)은 정신건강 낙인 수준(정신장애 병력이 있는 사람을 공직 출마에서 배제되어야 한다는 데 동의하는 인구의 비율로 측정함)과 역의 상관관계를 보였다. 따라서 낙인 수준이 높은 국가에서는 우울장애 보고를 꺼리기 때문에 우울장애의 유병률도 낮은 수준으로 나타난다⁸.

표 3.2. 낙인과 유병률 간의 관계는 복잡하지만 경우에 따라 낙인으로 인해 정신장애 유병률이 낮은 수준으로 보고될 수 있다

정신건강에 대한 낙인 지표와 정신질환 유병률 간의 상관관계, OECD 국가 중 유럽 9개국

| | 정신질환 병력이 있는 경우 업무에서 제외 | 치료에 대한 긍정적인 시선 | 정신건강을 다른 질병과 동일하게 생각 | 심리적 고통 유병률 (MHI-5) | 우울증 유병률 (PHQ-8) | 심리적 고통에 대한 응답 누락률 (MHI-5) |
|---------------------------|------------------------|----------------|----------------------|--------------------|-----------------|---------------------------|
| 정신질환 병력이 있는 경우 업무에서 제외 | 1 | | | | | |
| 치료에 대한 긍정적인 시선 | 0.11 | 1 | | | | |
| 정신건강을 다른 질병과 동일하게 생각 | 0.11 | 0.56 | 1 | | | |
| 심리적 고통 유병률(MHI-5) | 0.14 | -0.44 | -0.47 | 1 | | |
| 우울증 유병률(PHQ-8) | -0.83*** | -0.19 | -0.06 | 0.13 | 1 | |
| 심리적 고통 (MHI-5)에 대한 응답 누락률 | 0.17 | -0.6 | 0.11 | -0.001 | 0.03 | 1 |

주: 표는 목록별 상관관계를 표시한다. 낙인에 관한 3개의 문항에 대해 응답자는 진술에 어느 정도 동의하는지 표시한다. 첫 번째 문항의 경우 동의는 낙인을 의미하고 두 번째, 세 번째 문항의 경우 동의는 낙인이 없음을 의미한다. MHI-5 및 PHQ-8 척도에 대한 자세한 내용은 부속서 2.B를 참조한다. * 피어슨 상관계수(Pearson's correlation)는 p<0.10 수준에서 유의함을 나타내며 **는 p<0.05 수준에서, ***는 p<0.01 수준에서 유의함을 나타낸다. N = 9개국.

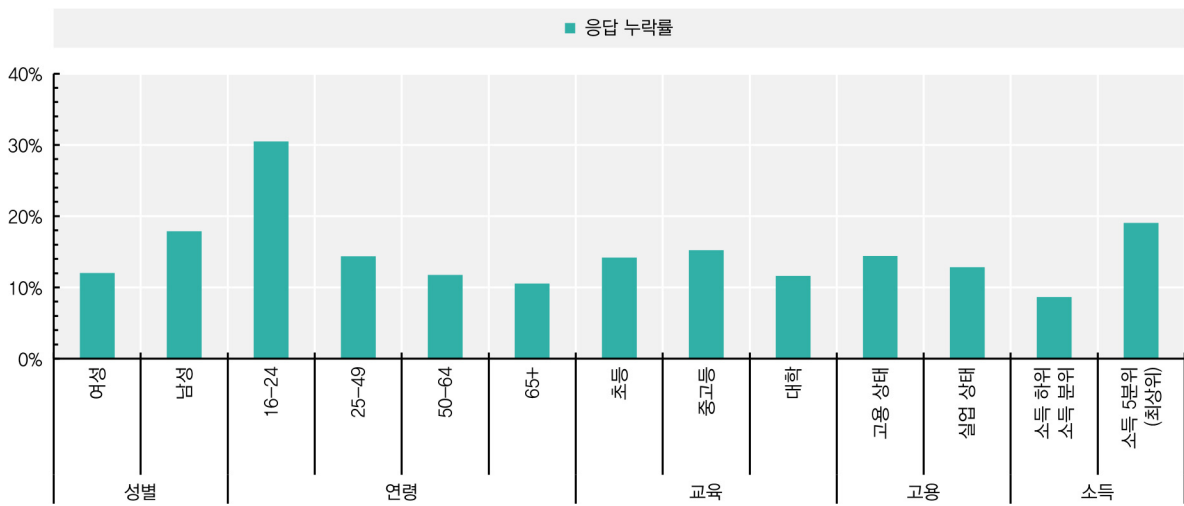
출처: 낙인 데이터는 OECD(2021^[57]), *Fitter Minds, Fitter Jobs: From Awareness to Change in Integrated Mental Health, Skills and Work Policies*, Mental Health and Work, OECD Publishing, Paris, <https://dx.doi.org/10.1787/a0815d0f-cn>에 게재된 입소스 설문조사에서 추출됨, MHI-5 데이터는 2018년 *유럽연합 소득 및 생활조건 통계(EU-SILC)* (n.d.^[65])(데이터베이스), <https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions>을 기반으로 OECD에서 계산함, PHQ-8는 *유럽건강면접조사(EHIS)* 데이터(n.d.^[66])(데이터베이스), [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:European_health_interview_survey_\(EHIS\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:European_health_interview_survey_(EHIS))를 기반으로 OECD에서 계산함.

무응답률에 의해 어떤 편향이 발생하는지 이해하려면 정신건강 문항에 응답하지 않기로 선택한 사람의 특성을 파악하는 것이 중요하다. 그림 3.2는 성별, 연령층, 교육 수준, 근로 상태 및 소득분위 등 다양한 사회인구통계학적 집단의 무응답률을 보여준다. 패널 A는 OECD 국가 중 유럽 26개국의 정신건강 관련 문항에 대한 무응답을 표시하고 패널 B는 영국의 무응답을 보여준다. 두 데이터 모두에서 여성, 교육 수준이 높은 집단, 고령층은 정신건강 관련 문항에 응답할 가능성이 높은 반면 남성, 청년층 및 교육 수준이 낮은 집단은 응답하지 않을 가능성이 높은 것으로 나타났다. 이러한 결과는 스웨덴의 정신건강에 대한 낙인을 설명하는 보고서의 내용과 일치한다. 보고서에 따르면 여성이 남성보다 정신질환을 앓고 있는 환자에 대해 긍정적인 감정을 가질 가능성이 더 높은 것으로 나타났으며, 청년층은 고령층보다 자신의 정신질환에 대해 타인에게 말하는 것이 어려울 것이라고 응답할 가능성이 더 높았다(Folkhälsomyndigheten, 2022^[67]).

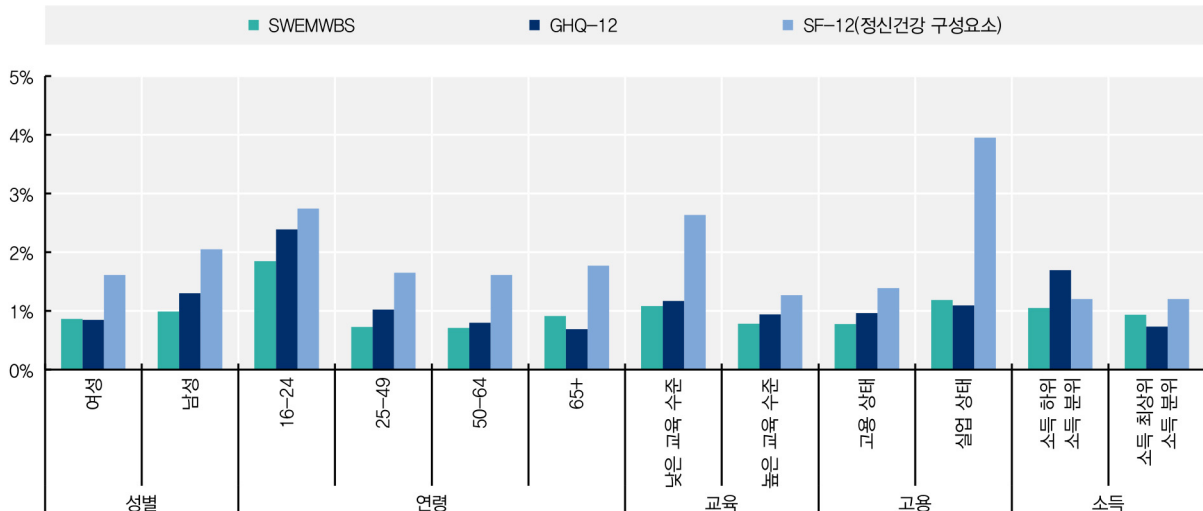
유럽 국가에서는 소득에 따라 정신건강 관련 문항에 대한 응답 여부에 명백한 차이가 나타났다. 소득 5분위(최상위)에 속하는 사람들은 해당 문항에 응답할 가능성이 낮았다. 반면 영국에서는 소득별 응답 차이가 나타나지 않았다. 호주 노인층을 대상으로 실시한 건강에 대한 종단조사의 무응답률에 대한 연구에 따르면 직업적 지위가 낮고 교육 수준이 낮은 사람들은 응답에 참여할 가능성이 낮지만(Jacomb 외, 2002^[68]), 우울이나 불안의 위험도도 응답 거부율에 영향을 미치지 않는 것으로 나타났다. 네덜란드 정신건강 설문 및 발생 연구-2(MENESIS-2)에서는 청년층의 무응답률이 높아 특정 정신장애에 대해 과소 보고된 것으로 나타났다(de Graaf, Have 및 Van Dorsselaer, 2010^[69]).

그림 3.2. 청년층, 남성 및 교육 수준이 낮은 집단은 정신건강 관련 문항에 응답할 가능성이 낮다

패널 A: 다양한 사회인구통계학적 집단별 정신건강 관련 문항(MHI-5)에 대한 무응답률, OECD 국가 중 유럽 26개국, 2013년




패널 B: 정신건강 관련 문항(SWEMWBS, GHQ-12 및 SF-12 정신건강 구성요소)에 대한 무응답률, 영국, 2019년



주: 특정 도구에 대한 자세한 내용은 부속서 2.B를 참조한다.

출처: 패널 A는 2013년 유럽연합 소득 및 생활조건 통계(EU-SILC) (n.d._[65]), (데이터베이스), <https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions>를 기반으로 OECD에서 계산함, 패널 B는 에식스대학교 사회경제연구소(2022_[70]), *Understanding Society: Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009* (데이터베이스). 15th Edition. UK Data Service. SN: 6614, [http://doi.org/10.5255/UKDA-SN-6614-16, from wave 10 only](http://doi.org/10.5255/UKDA-SN-6614-16_from_wave_10_only)(2018년 1월 - 2020년 5월)를 기반으로 OECD에서 계산함.

StatLink  <https://stat.link/1im7az>

특정 정신장애에 대한 이전 진단 또는 경험한 증상에 관한 데이터는 개인 병력을 공개하는 것을 꺼리는 심리와 함께 의료 서비스를 이용할 수 없기 때문에 인구 수준의 유병률을 과소 보고할 가능성이 높다(Hinshaw 및 Stier, 2008_[53]). 또한, 해당 데이터를 기반으로 도출된 정신질환 유병률은 광범위한 환자를 치료하고 진단할 수 있는 능력 등 국가 및 지역별 의료 시스템 특성에 따라 크게 영향을 받는다. 예를 들어 팬데믹 이전의 데이터에 따르면 정신건강 관련 의료 서비스의 이용을 희망하는 생산 연령 성인의 67%가 치료 접근성이 제한되어 있다고 보고한 것으로 나타났다(OECD, 2021_[71]). 캐나다의 한 조사에 따르면 캐나다 지역 건강조사에서 정신건강 의료 서비스 이용 현황을 퀘벡 정부(Régie de l'assurance maladie du Québec - RAMQ)의 의료 서비스 관련 행정적 데이터와 비교한 결과, 상당한 차이가 확인되었다. 퀘벡 정부의 레지스트리(RAMQ registry)에 등록된 정신건강 의료 서비스 이용자 중 75%가 캐나다 지역 건강조사(CCHS)에서 의료 서비스를 이용하고 있는 것으로 보고하지 않았으며, 이러한 격차는 고령층, 교육 수준이 낮은 집단 및 어린 자녀를 둔 학부모층에서 가장 두드러지게 나타났다(Drapeau, Boyer 및 Diallo, 2011_[72]). 호주를 대상으로 실시된 또 다른 연구에서는 자기보고 정신건강 정보(자기보고 처방약 사용 및 진단)를 정신장애 처방 기록과 대조하여 정신질환에 대해 과소 보고하는 정도를 조사한 결과, 정신질환은 낙인으로 인해 다른 건강 상태에 비해 과소 보고될 가능성이 매우 높다는 사실이 확인되었다(Bharadwaj, Pai 및 Suziedelyte, 2017_[73]).

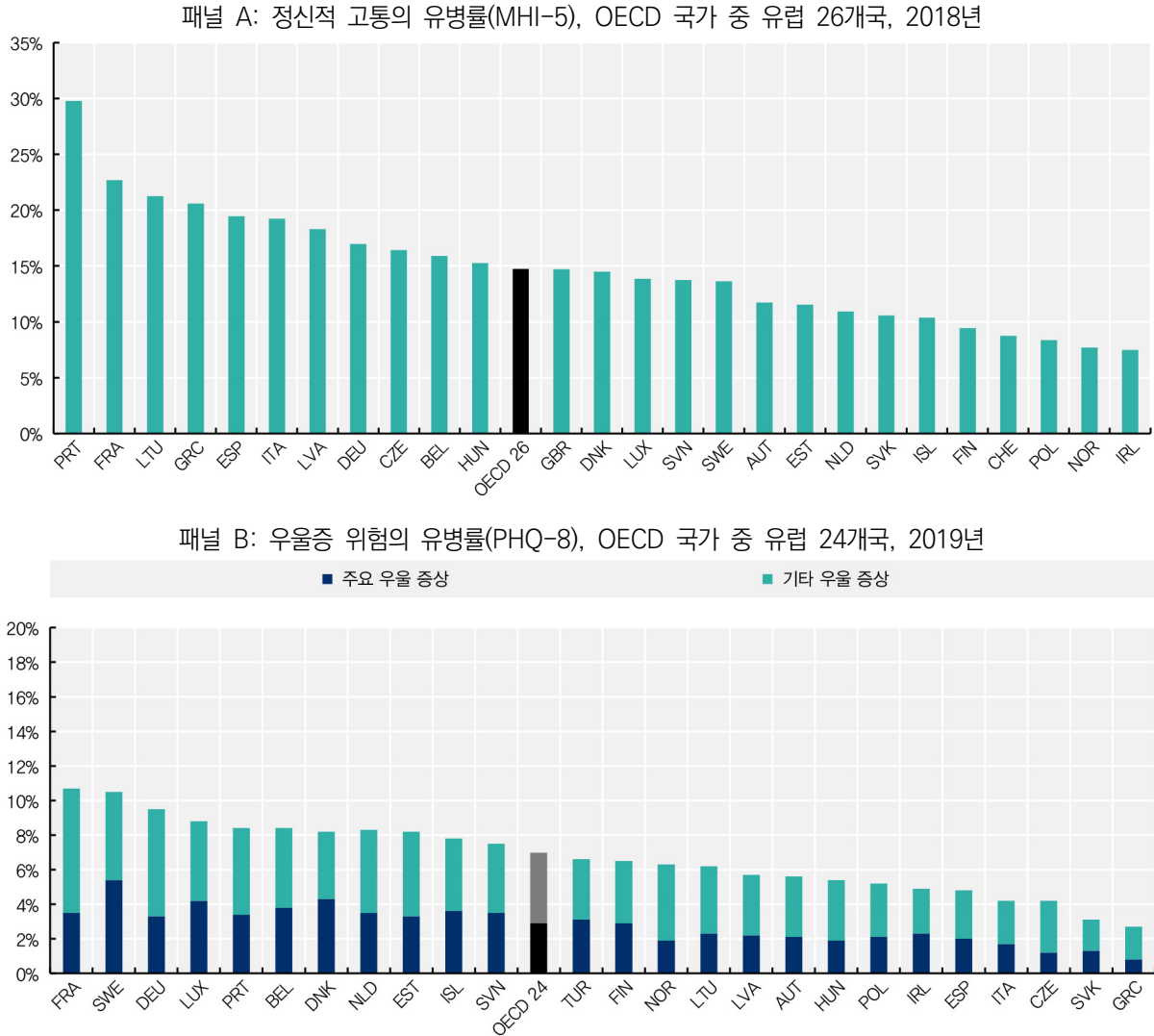
상자 3.6. 주요 메세지: 무응답 편향 및 결측치

- 기저 정신질환의 중증도가 심각할수록 설문조사 참여를 거부할 가능성이 높기 때문에 실제 정신질환 유병률이 과소평가될 수 있지만, 이는 설문조사 무응답률에 영향을 미치는 결정적인 원인은 아니다.
- 특정 정신건강 상태의 이전 진단 및 경험한 증상에 대한 자기보고 조사의 경우, 정신질환에 대한 낙인과 편견의 영향을 크게 받는다는 결정적인 근거가 있다.
- OECD 국가 중 유럽 국가의 분석에 따르면 청년층, 남성 및 교육 수준이 낮은 집단에서 정신건강에 관한 문항에 응답을 거부할 가능성이 높은 것으로 나타났다.

문화 및 사회인구통계학적 집단 간에 측정 신뢰도와 타당도가 일관되는가?


국민의 정신건강 증진을 목표로 관련 정책 및 조치를 수행하는 각국 정부는 국가별 정신건강 결과의 불평등을 이해하고 국가적 정신건강 동향(개선 또는 악화)에서 집단 간의 차이를 파악하기 위해서 고품질의 정보가 필요하다. 이를 통해 가장 도움이 필요한 대상에게 정책적 개입을 집중할 수 있다. 이러한 이유로 정신건강 지표는 연령 코호트, 성별, 인종 및 민족, 다양한 교육 및 소득 수준, 기타 사회경제적 지표를 비교할 수 있어야 한다. 그러나 비교 가능성을 보장하는 것은 간단하지 않다. 문화적 차이로 인해 정신건강에 대한 인식이 다르기 때문에 일부 집단에서는 정신건강과 관련된 문항에 응답(또는 정직하게 응답)할 가능성이 낮아질 수 있다. 이러한 문제는 국가 간 비교와 국가 내 비교에서 모두 나타난다⁹.

그림 3.3. 심리적 고통 및 우울 증상 위험 유병률의 경우 OECD 국가 중 유럽 국가 간에 100%까지 차이가 난다



주: 패널 A의 경우, 심리적 고통의 위험은 0점(최소)에서 100점(최대)까지 범위의 척도에서 점수가 52점 이상인 것으로 정의된다. 패널 B의 경우, 응답자가 PHQ-8의 처음 두 문항 중 한 문항에 ‘절반 이상’으로 응답하고 8개 항목 중 5개 이상이 ‘절반 이상’으로 보고되면 주요우울장애 위험군으로 간주한다. PHQ-8의 처음 두 문항 중 한 문항에 ‘절반 이상’으로 응답하고 8개 항목 중 총 2개-4개 항목이 ‘절반 이상’으로 보고되는 경우, ‘기타우울장애’ 위험군으로 간주한다 (Eurostat, n.d.^[74]). 개별 선별 도구에 대한 자세한 내용은 부속서 2.B를 참조한다.

출처: 패널 A는 2018년 *유럽연합 소득 및 생활조건 통계(EU-SILC)* (n.d.^[65]), (데이터베이스), <https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions>를 기반으로 OECD에서 계산함, 패널 B는 제3차 *유럽건강 면접조사(EHIS)* 데이터(n.d.^[66])(데이터베이스), [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:European_health_interview_survey_\(EHIS\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:European_health_interview_survey_(EHIS)).

StatLink  <https://stat.link/ocxvgt>

유럽 국가의 데이터에 따르면 심리적 고통과 우울 증상의 유병률에 큰 편차가 있음을 보여준다. 심리적 고통 유병률의 경우 포르투갈, 프랑스, 리투아니아에서의 유병률이 아일랜드, 노르웨이, 폴란드, 스웨덴보다 두 배 이상인 것으로 나타났다(그림 3.3, 패널 A). 이와 마찬가지로 프랑스, 스웨덴, 독일에서의 우울 증상 유병률은 그리스, 슬로바키아, 체코보다 두 배 이상인 것으로 나타났다(그림 3.3, 패널 B).

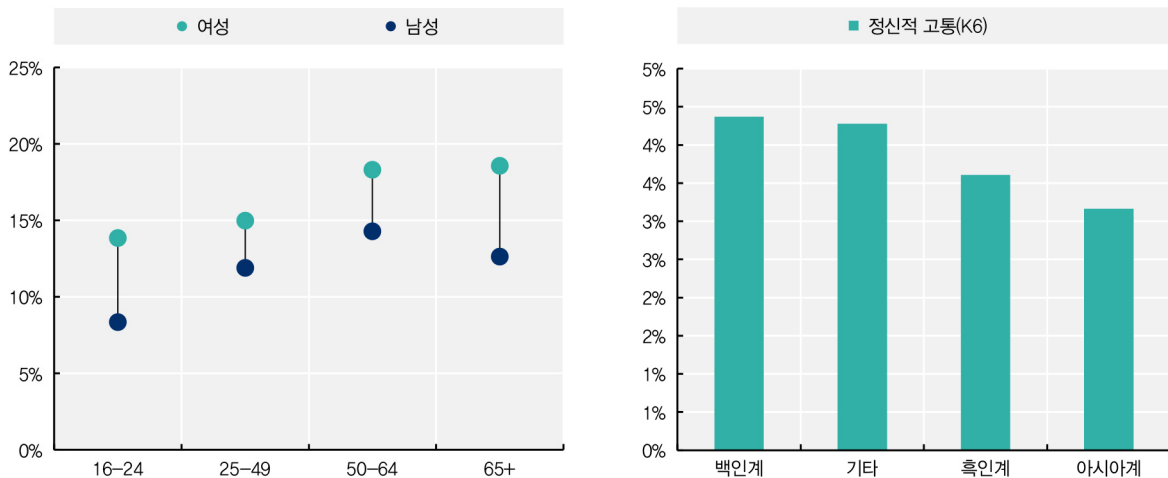
그러나 이러한 국가 간 편차의 원인이 나라별 인구의 근본적인 정신건강의 차이로 인한 것인지, 문화적 차이로 인해 선별 도구에 대한 응답 행동이 달라져 나타난 것인지 살펴볼 필요가 있다.

국가 간 편차는 국가별로 정신건강에 대한 낙인 수준이 다르다는 점에 일부 기인할 수 있는데, 전체 정신질환 유병률이 낮은 국가일수록 낙인 수준도 높은 것으로 나타났다(이전 섹션 및 표 3.2 참조).

국가 내 정신질환 유병률의 비교도 어려울 수 있다. 그림 3.4의 패널 A는 OECD 국가 중 유럽 26개국의 여성이 삶의 모든 단계에서 남성보다 더 높은 수준의 심리적 고통을 보고할 가능성이 높다는 것을 보여준다. 또한, 그림 3.4의 패널 B는 백인계 미국인이 다른 인종/민족 집단에 비해 심리적 고통의 수준이 높고, 아시아계 미국인은 심리적 고통의 수준이 가장 낮다는 것을 보여준다. 연구에 따르면 남성이 여성보다 자신의 증상을 최소화하려는 경향이 있기 때문에 자기보고 평가의 편향에는 성별에 따른 체계적인 차이가 있는 것으로 나타났다(Brown 외, 2018^[63]). 또한, 한 연구에 따르면 설문조사 동의서에 추후 후속 조치가 이루어질 수 있음을 명시한 경우, 여성은 아니지만 남성은 우울 증상을 더 적게 보고한 것으로 나타났다(Sigmon 외, 2005^[75]). 스웨덴의 정신건강 및 낙인에 대한 태도에 관한 조사에 따르면 여성이 남성보다 정신질환을 앓고 있는 대상에 대해 긍정적인 태도를 느끼는 것으로 보고할 가능성이 더 높다는 사실이 확인되었다(Folkhälsomyndigheten, 2022^[67]). 따라서 관찰된 편차 중 어느 정도가 실제 근본적인 정신건강의 차이에 기인한 것인지, 아니면 보고에서의 차이에 기인한 것인지 살펴볼 필요가 있다.

그림 3.4. 성별과 인종/민족별로 보고된 결과의 차이는 근본적인 정신건강의 차이 때문인가 아니면 측정 문제에서 비롯되었나?

패널 A: 성별 및 연령별 정신적 고통(MHI-5), OECD 국가 중 유럽 26개국, 2018년
패널 B: 인종/민족별 정신적 고통(K6), 미국, 2019년



주: 패널 A의 경우, 0점(최소)에서 100점(최고)까지의 척도에서 52점 이상을 심리적 고통의 위험군으로 정의한다. 패널 B의 경우, 0점(최소)에서 24점(최고)까지의 척도에서 13점 이상을 심리적 고통의 위험군으로 정의한다. 개별 선별 도구에 대한 자세한 내용은 부속서 2.B를 참조한다.

출처: 패널 A는 2018년 *유럽연합 소득 및 생활조건 통계(EU-SILC)* (n.d.^[65]), (데이터베이스), <https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions>를 기반으로 OECD에서 계산함, 패널 B는 미시간대학교 (2021^[76]), *Panel Study of Income Dynamics* (데이터베이스), <https://psidonline.isr.umich.edu/default.aspx> 데이터를 기반으로 OECD에서 계산함(2019년 데이터만 해당).

StatLink <https://stat.link/3yntk5>

측정 편향과 집단별 비교가능성에 대한 질문에 대답하기 위해 선별 도구를 임상적으로 검증하는 과정에서 설문조사가 주요 사회인구통계학적 특성에 따라 구조적 측정동일성을 갖는지 평가한다. 긍정적 정신건강에

대한 WHO-5 및 SWEMWBS 척도와 함께 우울증 및 불안 증상에 대한 선별 도구는 다양한 환경(예: 성별, 연령 코호트 및 인종/민족 집단)에서 가장 자주 검증된 도구이다.

특정 정신질환에 대한 선별 도구의 경우 PHQ-8과 GAD-7 모두 기본적인 성별 및 연령별 편향이 없는 것으로 밝혀졌다. PHQ-8 및 PHQ-9는 다양한 연령대, 문화 및 인종/민족 집단을 대상으로 여러 임상 환경에서 검증되었으며(El-Den 외, 2018_[8]), PHQ-2는 어린이 및 청소년을 대상으로 사용하도록 검증되었다(Richardson 외, 2010_[77]). 미국을 대상으로 실시된 PHQ-4에 대한 한 연구에 따르면 성별과 연령에 따른 구조적 측정동일성이 발견되지 않았으며, 이 결과는 인구 구조가 유사한 국가로 확장될 수 있지만 인구 구조가 상이한 국가의 경우 반드시 적용되는 것은 아니다(Sunderland 외, 2019_[33]).

또한 긍정적 정신건강 복합 척도는 연령과 성별에 따른 일반화 가능성에서도 우수한 성능을 보였다. WHO-5는 모든 연령대에서 구성 타당도가 우수한 것으로 나타났으며, 9세 이상의 어린이를 대상으로 사용하기에 적합한 것으로 간주되었다(Topp 외, 2015_[49]). MHC-SF는 성별, 연령 코호트 및 교육 수준 전반에 걸쳐 성능이 우수한 것으로 나타났다(Santini 외, 2020_[48]). WEMWBS의 경우 최초로 성인을 대상으로 타당화 되었지만, 이후 11세 이상의 청소년을 대상으로 사용하도록 타당화 되었다(Warwick Medical School, 2021_[78]). 14개의 항목으로 구성된 WEMWBS를 타당화하는 과정에서 두 항목에서 성별에 따른 편향성이 나타났다. 예를 들어 모든 정신적 웰빙 수준에서 남성은 여성보다 “나는 자신감이 높아졌다”는 항목에 긍정적으로 응답할 가능성이 더 높았다(Stewart-Brown 외, 2009_[45]). 이 두 항목은 7개의 항목으로 구성된 선별 도구(SWEMWBS)를 구성하는 과정에서 삭제되었다. SWEMWBS의 경우 성별, 결혼 여부 또는 가계 소득에 따른 응답률 차이가 나타나지 않는다(Tennant 외, 2007_[14]).

모든 도구에는 인종 집단 전반에 걸친 타당도에 대한 근거가 혼합되어 있으며, 대부분의 근거는 미국 또는 캐나다에서 도출되었다. 라틴계 및 영미계 인구를 대상으로 한 조사에서는 CES-D의 요인 구조가 문화 간 불변성을 가진다는 근거가 엇갈리고 있다(Crockett 외, 2005_[79], Posner 외, 2001_[80]). 한 연구에 따르면 고령의 히스패닉/라틴계 및 흑인계 응답자를 대상으로 CES-D를 실시할 경우, 항목 수준에서 수정이 필요하다는 사실이 확인되었다(El-Den 외, 2018_[8]). 다른 연구에서도 아시아계 미국인과 아르메니아계 미국인은 영미계 미국인에 비해 요인 구조가 다르고 우울 증상이 더 높으며, 긍정적 정서에 관한 항목에 과도하게 응답하는 경향이 있는 것으로 나타났다(Iwata 및 Buka, 2002_[81], Demirchyan, Petrosyan 및 Thompson, 2011_[82]). 영국에서 다양한 인종 및 민족 집단을 대상으로 GHQ-12를 시행한 연구에서는 집단별 차이에 대한 시사적인 근거가 일부 발견되어 추가 연구가 필요하다(Bowe, 2017_[83]). 한국계 미국인과 영미계 미국인의 고령층을 비교한 연구에서는 문화적 요인이 우울증 척도의 진단 정확도에 유의한 영향을 미칠 수 있으며, 잠재적으로 인구집단에 따라 준거점수를 다르게 사용할 수 있음이 확인되었다(Lee 외, 2010_[84]). 또 다른 연구에서는 범불안장애 증상이 심각한 흑인계/아프리카계 미국인 응답자의 GAD-7 점수가 유사한 수준의 범불안장애 증상이 있는 기타 인종 및 민족의 응답자 점수보다 낮은 것으로 나타났다(Parkerson 외, 2015_[85], Sunderland 외, 2019_[33]).

자기보고 정신건강 측정도 인종/민족 정체성에 따른 편향성에 취약한 것으로 나타났다. 미국의 연구에 따르면 인종에 따라 자기보고 정신건강(SRMH)과 정신건강 상태의 범위 간의 관계가 조정되는 것으로 나타났다. 예를 들어 흑인계 및 히스패닉/라틴계 미국인은 백인 미국인보다 SRMH를 우수한 수준으로 보고할 가능성이 더 높으며, SRMH와 의료 서비스 이용 현황 간의 연관성이 더 낮은 것으로 나타났다. 캐나다의 한 연구에 따르면 사회경제적 상태를 통제한 후에도 아시아계 정체성은 열악한 SRMH와 관련이 있는 것으로 나타났다(Ahmad 외, 2014_[51]).

다수의 선별 도구가 여러 언어로 번역되어 전 세계 설문조사에 사용되고 있다. WHO-5, K10, MHI-5, GAD-7 및 WEMWBS는 여러 언어로 번역되었다(Sunderland 외, 2019_[33]). 예를 들어 WHO-5는 30개 이상의 언어로 번역되었으며, 6개 대륙의 설문조사에서 시행되었다(Topp 외, 2015_[49])¹⁰. WEMWBS는 50개국에서

사용되었으며, 36개 언어로 번역되었다(Stewart-Brown, 2021^[16], Warwick Medical School, 2021^[78]). MHC-SF에 대한 심리측정적 속성 평가도 많은 국가에서 수행되었는데(Petrillo 외, 2015^[47], Joshanloo 외, 2013^[86], Guo 외, 2015^[46]), 국가 간 번역에 대한 비교에서 높은 변동성을 보여준다. 그중 일부는 잠재적인 정신건강의 차이가 아닌 측정 문제에서 비롯된 것일 수 있다(Santini 외, 2020^[48]).

문화적 차이는 정신건강에 대한 통일된 정의와 설명을 확립하는 데 상당한 어려움을 야기하며, 국가 간 비교 가능성에도 위협 요소로 작용한다(상자 3.4 참조). 문화 간 타당도 검증은 특정 문화권에서 개발된 정신건강 관련 척도가 다른 문화에서도 적용 가능하고 의미가 있으며, 이에 따라 동등한 성능을 수행하는지 여부를 나타낸다(Huang 및 Wong, 2014^[87]). 가장 널리 사용되는 정신건강 척도는 영어를 사용하는 고소득 서구권 국가 인구(예: 북미, 유럽, 호주)에서 개발되고 검증되었기 때문에 정신장애 및 증상에 대한 서구적 이해를 전제로 한다(Sunderland 외, 2019^[33]). 이에 따라 다른 인구집단에 대한 적용 가능성에 의문이 제기될 수 있다. 예를 들어 난민의 정신건강 상태에 대한 연구 183건을 검토한 결과, 연구 결과의 78%가 난민 인구를 대상으로 특화하여 개발되거나 검증되지 않은 도구를 기반으로 수행된 것으로 나타났다(Hollifield 외, 2002^[88]).

다양한 도구에 대한 문화 간 검증에 대한 근거가 엇갈리고 있다. WEMWBS는 17개의 언어와 지역 집단 뿐만 아니라 영국 내 소수 집단에 대해서도 검증되었다(Warwick Medical School, 2021^[78]). PHQ는 여러 환경에서 검증되었고 강력한 선별 도구 중 하나로 간주되지만, 한 연구에 따르면 중소득 또는 저소득 국가에 적용했을 경우 임상 표본이 아닌 학생 표본에 한해서 성능이 우수하여 비교적 정신건강에 대한 이해도가 높은 환경에서만 사용하도록 제안된다(El-Den 외, 2018^[8], Ali, Ryan 및 De Silva, 2016^[32]). 마찬가지로 점수 체계, 즉 선별 도구에서 특정 정신 문제의 위험을 나타내는 점수를 결정하는 과정은 최초 임상 연구가 수행된 미국의 일반 인구를 기준으로 보정되는 경우가 많다. 케슬러 척도의 점수 체계는 일반 인구 분포의 90%~99% 백분위수에서 최대 정확도를 구하도록 설계되었는데, 이는 특정 연도에 미국 인구의 6%~10%가 심각한 정신질환을 앓고 있다는 미국의 역학 조사 결과에 기인한다. 따라서 이러한 점수 체계는 다른 구조를 가진 인구집단에는 적합하지 않을 수 있다(Kessler 외, 2002^[89]). 또 다른 예로는 SF-12의 정신건강 구성요소의 경우 일반적으로 각 응답 범주에 대해 미국에서 도출한 항목별 가중치를 사용하여 점수를 산출한다(Ware 외, 2002^[90]). 유럽과 호주에서 국제 비교를 통해 이러한 항목별 가중치의 적용이 적절한 것으로 밝혀졌으나(Vilagut 외, 2013^[91]) 해당 결과가 다른 지역으로 반드시 확장되는 것은 아니다.

상향식 및 개방형 접근 방식을 사용하여 개발되거나 현지 적응도가 높은 문화적으로 특화된 정신건강 척도에 대한 연구가 필요하다. 다양한 문화와 민족에 걸쳐 기존 척도에 대한 추가적인 검증도 반드시 필요하다(Sunderland 외, 2019^[33]). 아울러 심리측정적 모형과 컴퓨터를 사용한 통계의 발전으로 선별 도구의 관리 및 채점 방식이 새롭게 개발되어 문화 간 분석이 용이해졌다¹¹. 그러나 이러한 차이의 규모를 전후 사정에 맞게 파악하는 것이 중요하다. 150개국을 대상으로 진행된 삶의 만족도를 포함한 긍정적 정신건강 측정의 국가 간 차이에 대한 Gallup World Poll의 데이터를 사용한 한 연구에 따르면 국가 간 차이 중 문화적 차이는 20%에 불과한 것으로 나타났다. 20%에는 다양한 문화가 결과에 미치는 영향과 잠재적인 측정 편향성이 모두 포함되며, 이는 소득, 교육, 고용과 같은 객관적인 조건의 영향에 비하면 미미한 수준이다(Exton, Smith 및 Vandendriessche, 2015^[92], OECD, 2013^[11]). 이러한 객관적인 삶의 조건이 정신건강에 미치는 영향도 문화적 차이의 영향보다 클 가능성이 높다. 인구집단 전반에 걸쳐 정신건강 도구를 더 잘 설계하고 검증하는 것이 중요하다는 사실을 부정하는 것은 아니지만, 정신건강 지표가 정책에 유익하고 유용한 정보를 제공한다는 점을 기억할 필요가 있다.

상자 3.7. 주요 메시지: 집단 간 정확성

- 정신건강에 대한 태도의 차이는 성별, 연령, 인종/민족 정체성뿐만 아니라 국가 간에도 보고에 차이를 초래할 수 있다.
- 낙인과 차별적 견해에 관한 설문조사에 따르면 정신건강에 대한 태도는 연령과 성별에 따라 차이가 있는 것으로 나타났다.
- 선별 도구는 타당화 과정에서 연령, 성별, 인종/민족 집단에 따른 편향성 유무를 검증한다. 특정 정신건강 상태에 대한 선별 도구와 긍정적 정신건강 복합 척도는 대부분의 경우 연령과 성별에 대해 우수한 성능을 보이지만 인종/민족에 대한 근거는 엇갈리고 있다. 일반적인 정신건강에 대한 자기 보고 문항의 성능에 대해서는 향후 추가적인 연구가 필요하다.
- 설문조사 항목의 적합성을 보장하기 위해 현지 인구집단을 대상으로 타당화 되어야만 한다. 한 지역이나 인구집단에서 수행된 검증은 다른 상황에 적용되지 않을 수 있다.

시간이 경과함에 따라 측정치가 비교 가능한가?

정책 입안자의 주요 목표는 시간 경과에 따른 전반적인 정신건강 추세를 파악하는 것이다. 인구 수준에서 정신건강이 개선 또는 악화되고 있는지 혹은 정책적 개입이 정신건강 결과에 가시적인 변화를 초래하는지 등을 파악한다. 따라서 선택한 지표의 정확성이 횡단적으로 뿐만 아니라 시간이 경과함에 따라 유지되는 것이 중요하다. 시간이 경과함에 따라 정신건강을 측정하는 데는 두 가지 복잡한 사안이 있다. (1) 정신건강에 대한 행동 및 태도 변화로 인해 응답 행동이 달라질 수 있고, (2) 다수의 선별 도구가 횡단 연구에서 임상 진단에 대해 타당화 되었기 때문에 시간 경과에 따른 변화에 민감하다는 충분한 근거를 제공하지 못할 수 있다는 사실이다.

정신건강에 대한 태도는 수년에 걸쳐 변화해 왔으며 낙인과 편견이 여전히 남아있지만 이를 줄이는 데 진전이 있었다. 최근 몇 년 동안 OECD 각국 정부는 특히 학교와 교육 기관을 중심으로 정신질환에 대한 낙인을 없애기 위해 공공 정보 캠페인을 추진해왔다. 코로나19 팬데믹 이전에도 OECD 12개국은 정신건강에 대한 이해도 향상을 위한 국가적 캠페인을 벌였고 5개국에서는 지역 캠페인을 추진했다(OECD, 2021^[71]). 초기에는 캠페인의 영향에 대한 근거가 엇갈렸다. 일부 연구에서는 특히 장기적으로 정신건강 상태에 대한 낙인이 거의 또는 전혀 감소하지 않는 것으로 나타난 반면(Deady 외, 2020^[93], Walsh 및 Foster, 2021^[94]), 다른 연구에서는 정신과 응급실 방문과 같은 의료 서비스의 이용이 증가한 것으로 나타났다(Cheng 외, 2016^[95]). 영국의 한 연구에 따르면 정신건강 캠페인에 노출된 청년층에서 이러한 현상이 증가했을 수 있으며, 이는 기존의 감정에 대한 새로운 인식에서 비롯된 것이 아니라 정신질환에 대한 정보 증가로 인한 인과적 결과물이다(Harvey, n.d.^[96]). 이러한 맥락에서 다른 초기 연구에 따르면 인식개선 캠페인은 개인이 자신의 감정과 정서를 경증 또는 중증의 정신적 고통 징후로 분류하도록 유도하여 인식과 행동을 변화시켜 실제 증상이 악화될 수 있다고 가정한다(Foulkes 및 Andrews, 2023^[97]).

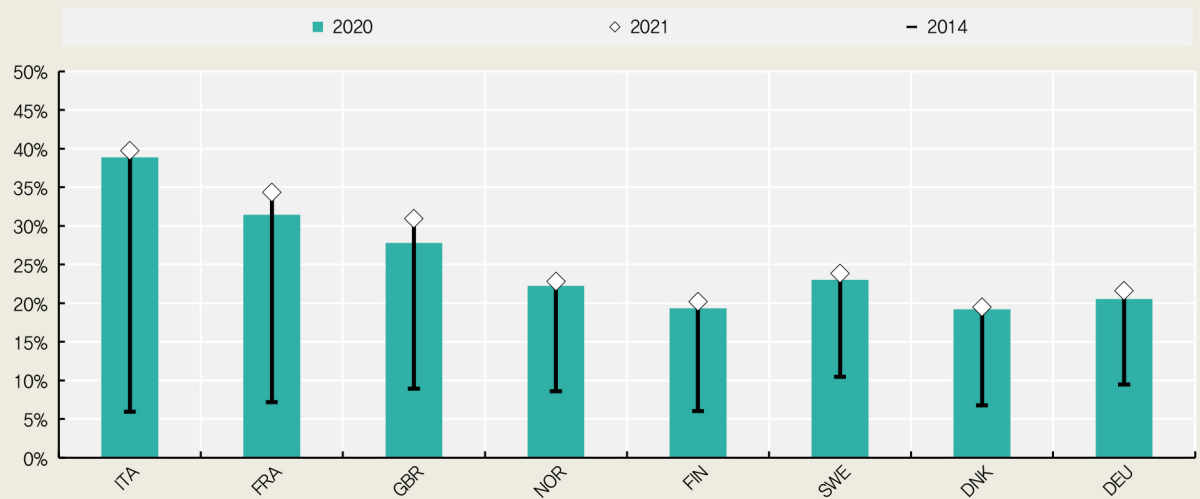
낙인 방지 캠페인이 실제 의도한 목표대로 영향을 미친다면 정신건강 문제에 대한 일반 인구의 태도가 변화하고 보통의 사람들이 자신의 정신건강에 대해 공개적이고 정직하게 말하는 것을 더욱 편하게 느낄 수 있다. 이는 시간이 경과함에 따라 정신질환 유병률의 추정치를 왜곡할 수 있다. 10년 전 일반 인구가 지난 2주 동안 ‘울적함, 우울함 또는 절망감’을 얼마나 자주 느꼈는지에 대한 문항에 정직하게 응답하는 것을 불편하게 느꼈다면, 자신의 진짜 감정을 속이고 응답하지 않는 비율이 오늘날보다 더 높았을 것이다. 따라서 이러한 태도 변화로 인해 심리적 고통의 유병률에 대한 보고가 증가된 것처럼 보여질 수 있다.

상자 3.8. 팬데믹 기간 동안의 정신건강의 변화

코로나19 팬데믹 기간과 그 이후 대부분의 OECD 국가에서 정신건강이 악화되었으며, 일부 국가에서는 우울증과 불안 증상의 비율이 두 배로 증가한 것으로 나타났다(OECD, 2021^[98], OECD, 2021^[57]). 실제로 팬데믹 이전에 수집된 기준 데이터가 유사한 OECD 국가 중 유럽 8개국의 경우 우울증 위험군 비율이 크게 증가했으며, 이탈리아와 프랑스에서도 20% 포인트 이상 증가한 것으로 나타났다(그림 3.5) (OECD, 2021^[99]). 2020년 1월부터 2021년 1월까지의 데이터를 조사한 한 연구에 따르면 2020년에는 팬데믹 발발 이전에 비해 불안과 우울증 증상을 경험한 사람의 비율이 각각 28%와 26% 더 높은 것으로 추정된다(OECD, 2021^[57])¹². 유럽 국가의 종단적 및 횡단적 연구에서 WHO-5(정서 기반 측정) 및 SWEMWBS 또는 MHC-SF(정서, 행복, 사회적 관계 및 삶의 평가 측면 결합)을 통해 측정된 긍정적 정신건강이 팬데믹 기간에 크게 악화된 것으로 나타났다(Thygesen 외, 2021^[100], Vistisen 외, 2022^[101], Eurofound, 2021^[102], Vinko 외, 2022^[103]).

정신질환 증상의 유병률이 증가한 현상에 대해 전반적으로 의견이 일치하지만, 이러한 증가가 일시적인 것인지 아니면 정신건강이 비교적 빠르게 팬데믹 이전 수준으로 회복할 것인지는 아직 미지수이다. 2021년 중반 현재 전반적인 정신건강은 팬데믹 이전 수준으로 회복되지 않았으나 일부 OECD 회원국에서는 회복 조짐을 보이고 있다(OECD, 2021^[57], OECD, 2021^[99]). 그러나 여전히 청년층과 같이 특히 부정적인 영향을 받은 특정 인구집단은 많은 어려움에 직면해 있다(OECD, 2021^[99]).

그림 3.5. 팬데믹 첫해에 OECD 국가 중 유럽 8개국에서 우울증 증상이 크게 증가했다
우울증 위험군 비율, 2020년 및 2021년(2014년 대비)



주: 2020년 및 2021년 데이터는 2014년 데이터와 다른 데이터 소스에서 추출된 것으로 개별 국가의 수치 증가를 해석할 때 주의해야 한다. 두 데이터 소스 모두 우울증 위험의 척도로 PHQ-2를 사용한다. 2020년 및 2021년 데이터는 YouGov 코로나19 행동 추적기(YouGov COVID-19 behaviour tracker)에서 추출되었다. 2020년 통합 평균은 4월부터 12월까지, 2021년 통합 평균은 1월부터 6월까지의 데이터이다. 기준 데이터는 2014년 제2차 유럽건강면접조사(EHIS)에서 추출된 것이다. 개별 선별 도구에 대한 자세한 내용은 부속서 2.B를 참조한다.

출처: OECD(2021^[99]), COVID-10 and Well-being: Life in the Pandemic, OECD Publishing, Paris, <https://dx.doi.org/10.1787/1e1ecb53-en>.

StatLink <https://stat.link/szdacx>

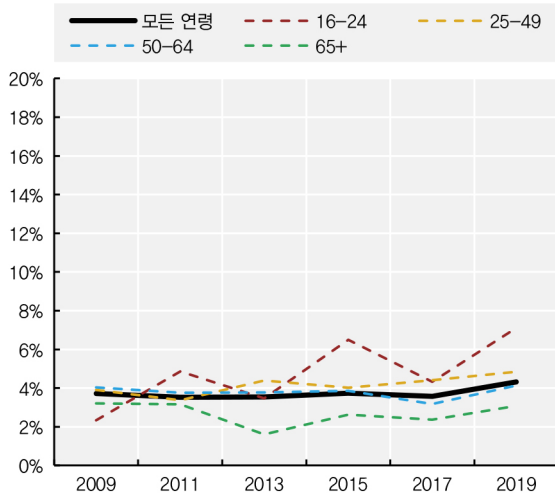
그림 3.6에서는 이 가설을 반증하는 몇 가지 근거를 제시하고 있다. 20개국 이상의 OECD 유럽 국가 및 미국의 팬데믹 이전 데이터에 따르면 금융 위기 이후부터 코로나19 이전까지 지난 몇 년 동안 정신건강이 개선되거나 안정된 것으로 나타났다(불안 및 우울증 증상의 유병률은 팬데믹이 시작된 2020년에 급격히 증가함, 상자 3.8 참조). 2009년부터 2019년까지 미국의 심리적 고통 유병률(K6 선별 도구를 사용하여 격년으로 측정)은 10년 동안 약 4%(패널 A)에서 대체로 안정적으로 유지되었다. 사회인구통계학적 요인을 통제하지는 않지만, 이는 높은 유병률로 이어지는 응답 행동이 크게 변화하는 등 인식 변화에 대한 우려가 지속되지 않을 수 있음을 시사한다. 즉 16세~24세의 청년층에서 유병률이 더 높거나 잠재적으로 증가할 수 있다는 몇 가지 근거가 있으며, 이는 변화하는 상황(사회정치적, 경제적, 기후 관련), 정신건강에 대한 청년층의 태도 변화(조사원에게 공개적으로 말하려는 태도), 청년층 코호트의 작은 표본 크기(데이터에서 더 많은 노이즈를 초래) 등이 복합적으로 반영된 것일 수 있다. OECD 국가 중 유럽 26개국에서 2013년~2018년에 심리적 고통이 감소했는데, 이는 행동 변화로 인해 응답자가 자신의 정신건강이 좋지 않다고 솔직하게 응답할 가능성이 높아졌다면 예상할 수 없는 결과이다(패널 C). 패널 D에서도 이와 유사한 결과가 나타났는데, 2011년~2016년에 OECD 국가 중 유럽 24개국에서 심리적 안정감이 증가한 것으로 나타났다. 반대로 영국의 데이터(패널 B)는 인구 수준의 정신건강이 악화(SF-12의 정신건강 구성요소로 측정)된 반면, 긍정적 정신건강(SWEMWBS)과 일반적인 정신장애 위험군의 비율(GHQ-12)은 다소 안정적인 상태를 유지하는 것으로 나타났다.

일부 인구 조사에서 시간이 경과함에 따라 비교적 안정적인 정신건강 유병률이 나타난 원인 중 한 가지는 해당 측정 도구가 변화에 대한 민감도가 낮아져 기인할 수 있다. 많은 정신건강 선별 도구는 횡단적 임상 표본에서 검증되었다. 따라서 종단적 데이터로만 평가할 수 있는 시간 경과에 따른 변화에 대한 민감도가 검증되지 않았음을 주의해야 한다(Ahmad 외, 2014^[51], Tennant 외, 2007^[14], Moriarty, Zack 및 Kobau, 2003^[21], Spitzer 외, 2006^[11]). 그러나 예외도 존재한다. 일부 연구에 따르면 PHQ-8과 PHQ-2는 시간 경과에 따른 변화에 민감하게 반응하는 것으로 나타났다(Löwe 외, 2010^[13]). 긍정적 정신건강의 경우, WEMWBS와 단축형 버전인 SWEMWBS는 집단과 개인에서 모두 시간 경과에 따른 변화에 민감한 것으로 나타났다. 연구자에 따르면 WEMWBS 척도에서 +/- 3점, SWEMWBS 척도에서 +/- 1점~3점이 유의미한 변화를 나타낸다(NHS Health Scotland, 2016^[15], Shah 외, 2018^[104]).

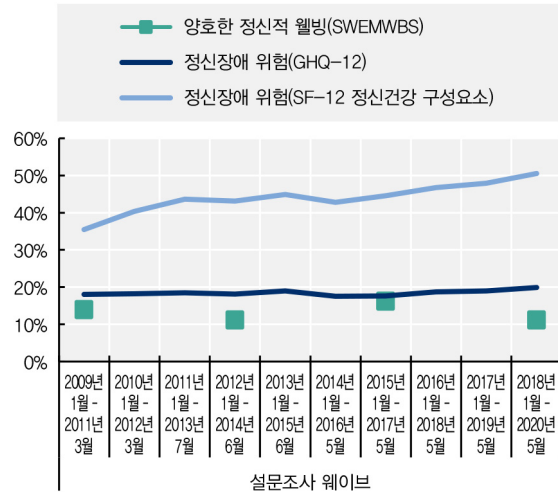
시간 경과에 따른 개인의 특성 또는 환경의 변화로 인한 측정 편향을 식별하는 한 가지 접근 방식은 종단적 데이터를 사용하는 것이다. Ploubidis 외 (2019^[105]) 연구는 영국에서 대표적인 두 개의 국가 설문조사를 사용하여 20년에 걸쳐 연령 코호트를 추적했다. 잠재변수를 측정하는 일반적인 측정 모델링 프레임워크를 사용하여 불쾌감(Malaise) 검사(심리적 고통을 측정하는 9개 항목으로 구성된 설문조사)의 문항에 대한 응답자의 답변이 시간 경과에 영향을 받는지를 검증한 결과, 연령 효과, 설문조사 설계, 기간 효과 및 코호트별 효과의 형태로 편향성이 존재한다는 근거를 거의 발견할 수 없었다.

그림 3.6. 2019년까지 OECD 유럽 국가에서는 정신건강이 다소 개선되었으며 미국은 정신건강에 대한 인식이 개선되었음에도 대체로 안정적인 수준을 유지했다

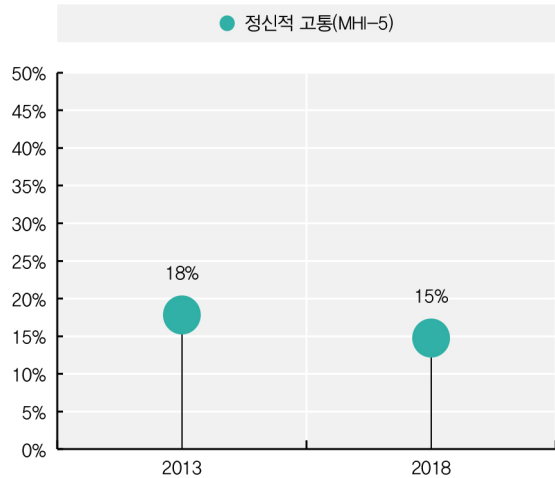
패널 A: 정신적 고통 위험군(K6)의 연령대별 비율, 미국, 2009년~2019년



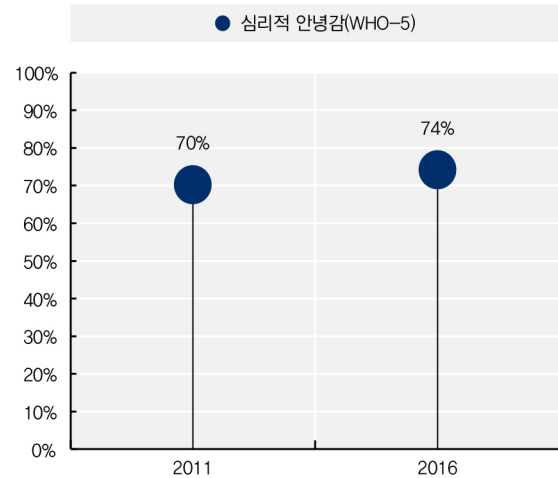
패널 B: 정신건강이 양호한 인구(SWEMWBS) 또는 정신장애 위험군의 비율(GHQ-12 및 SF-12의 정신건강 구성요소), 영국, 2009년~2020년



패널 C: 정신적 고통 위험군(MHI-5)의 비율, OECD 국가 중 유럽 26개국, 2013년 및 2018년



패널 D: 심리적 안녕감 인구(WHO-5)의 비율, OECD 국가 중 유럽 24개국, 2011년 및 2016년



주: 패널 A의 경우 각 선별 도구에 대한 채점 정보는 다음과 같다. 0점(최소)에서 24점(최대)까지의 척도에서 13점 이상인 경우 심리적 고통의 위험군으로 정의한다. 패널 B의 경우 변환된 SF-12 정신건강 구성요소 복합 척도에서 점수가 50점 이하인 경우 정신질환 위험 측면에서 0점은 최악의 정신건강 상태, 100점은 최상의 정신건강 상태를 나타낸다. Woodhead 외(2012_[106])에서 사용된 바와 같이 GHQ-12에서 점수가 4점 이상인 경우 일반적인 정신장애(CMD) 가능성, 양호한 정신건강은 SWEMWBS 점수가 표본 평균보다 1 표준편차 이상 높은 것으로 정의한다. 패널 C는 0점(최소)에서 100점(최대)까지의 척도에서 52점 이상인 경우 심리적 고통 위험을, 패널 D는 0점(최악)에서 24점(최고)까지의 척도에서 14점 이상인 경우 심리적 안녕감을 나타낸다. 자세한 내용은 부속서 2.B를 참조한다.

출처: 패널 A는 미시간대학교(2021_[76]), *Panel Study of Income Dynamics* (데이터베이스), <https://psidonline.isr.umich.edu/default.aspx>를 기반으로 OECD에서 계산함, 패널 B는 에식스대학교 사회경제연구소(2022_[70]), *Understanding Society: Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009* (데이터베이스), 15th Edition. UK Data Service. SN: 6614, <http://doi.org/10.5255/UKDA-SN-6614-16>를 기반으로 OECD에서 계산함, 패널 C는 2013년 및 2018년 *유럽연합 소득 및 생활조건 통계(EU-SILC)* (n.d._[65]), (데이터베이스), <https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions>를 기반으로 OECD에서 계산함, 패널 D는 2011년 및 2016년 *유럽 삶의 질 조사*(n.d._[107]), (데이터베이스), <https://www.eurofound.europa.eu/surveys/european-quality-of-life-surveys>를 기반으로 OECD에서 계산함.

StatLink <https://stat.link/y5vgfp>

상자 3.9. 주요 메세지: 시간 경과에 따른 정확성

- 정신건강 상태에 대한 태도의 변화와 정신건강 관련 주제를 공개적으로 논의하는 데 편안함을 느끼는 것은 시간 경과에 따른 유병률을 비교할 경우 측정 편향으로 이어질 수 있다.
- 낙인 방지 캠페인에 대한 영향 평가에서 도출된 근거는 한정적이지만 팬데믹 이전에 국가별 추세 데이터에서 도출된 근거에 따르면 낙인의 감소가 심리적 고통에 대한 보고를 확실히 증가시킨다고 나타나지는 않았다.
- PHQ-8 및 (S)WEMWBS를 포함한 일부 정신건강 선별 도구는 종단적 연구에서 시간 경과에 따른 변화에 민감한 것으로 나타났지만 다른 도구는 민감도에 대한 분석이 이루어지지 않았다.
- 정신건강 결과에 대한 연령, 기간 또는 코호트별 영향의 형태로 편향성이 존재하는지에 대해 향후 추가적인 연구가 수행되어야 한다.

데이터 수집

데이터 수집의 실용성은 응답자의 행동에 중요한 영향을 미칠 수 있으며, 응답자가 조사원과 상호 작용하고 문항에 답하는 데 있어 편안함과 용이성에 영향을 미칠 수 있다. 문항에 구성된 진술이 긍정적이거나 부정적인 것과는 관계없이 데이터 수집 방식(조사원 또는 자기기입식)에 따라 최종 결과물의 품질이 결정된다. 특히 자살 및 자살 사고에 대한 선별 도구의 경우, 정신건강 관련 문항의 민감한 특성으로 인해 응답자와 조사원 모두의 안전과 웰빙을 보장하기 위한 추가 프로토콜을 마련해야 한다.

문항의 문구가 응답자의 태도와 응답 행동에 어떤 영향을 미치는가?

문항의 순서와 문항의 구성 방식에 따라 응답자가 특정 방식으로 답변하도록 유도할 수 있다. 주관적 웰빙에 대한 OECD 연구에 따르면 생애 평가와 정서에 관한 문항의 질문 순서가 미치는 영향이 상당히 클 수 있으므로 주관적 웰빙에 관한 문항은 다른 모듈의 영향력을 최소화하기 위해 설문조사 초반부에 배치되어야 한다(OECD, 2013_[11]).

정신건강 관련 문항의 경우 응답자에게 화를 유발하여 윤리적 문제를 초래할 수 있다는 일부 근거가 존재한다. 일부 연구에 따르면 응답하기 곤란한 문항이 포함된 설문조사에 참여하거나 인생에서 고통스러운 사건에 초점을 맞춘 문항에 응답하는 것은 응답자의 스트레스를 증가시키고 기분을 악화시킬 수 있다(Labott 외, 2013_[108]). 특히 이미 심리적 고통의 위협에 처한 인구집단에서 이러한 특징이 더욱 두드러지게 나타났다. 그러나 정신건강 설문조사가 응답자의 기분에 미치는 영향에 대한 다른 연구에서는 유의미한 근거를 발견하지 못했다(Jorm 외, 1994_[109], Jacomb 외, 1999_[110]). 응답에 괴로움을 느낀다고 답한 피면담자 중 소수는 젊은 여성과 사회적 지원이 결여된 사람일 가능성이 높았다(Jacomb 외, 1999_[110]).

특정 정신건강 선별 도구 또는 복합 척도 내에서 긍정적이거나 부정적인 관점에서 문항을 구성하면 응답에 영향을 미칠 수 있다. 부정적인 문항으로만 구성된 도구(예: PHQ-8, CES-D, K6), 긍정적인 문항으로만 구성된 도구(예: (S)WEMWBS, WHO-5), 두 가지를 혼합하여 사용하는 도구(예: GHQ-12, MHI -5) 등이 있다. 예를 들어 부정적인 문항으로 구성된 도구의 경우 얼마나 자주 낙담하고 우울한지에 대해 질문하는 반면, 긍정적인 문항으로 구성된 도구에서는 얼마나 자주 기분이 밝고 경쾌했는지에 대해 질문할 수 있다. 응답자는 ‘항상’ 우울함을 느낀다고 응답하는 것보다 기분이 밝았던 적이 ‘거의 없음’이라고 응답하는 데 더욱 편안함을 느낄 수 있다. 이는 영국가구종단연구(UKHLS) 설문조사 데이터를 기반으로 하는 표 3.3에

설명되어 있다. 세 가지 정신건강 선별 조사에서 동일한 표본의 응답자에게 동일한 질문을 한 결과 문항의 유형과 내용이 중복되는 부분이 있으며 한 쌍의 문항이 표의 상단에 표시되어 있다. 응답의 상관관계는 문항의 진술이 유사한 어조(긍정적 또는 부정적)로 구성된 도구에 포함된 문항(예: SF-12의 정신건강 구성 요소에서 낙담하고 우울한 느낌과 GHQ-12의 불행하고 우울한 느낌)에서 가장 높고 다른 어조로 구성된 도구에 포함된 항목(예: GHQ-12의 문제에 직면할 수 있었음과 SWEMWBS의 문제에 잘 대처함)에서 가장 낮게 나타났다.

영국은 정신건강 관련 의료 서비스를 이용하는 사람의 경우, 부정적인 감정보다는 긍정적인 감정에 초점을 맞춘 설문조사 도구를 선호하는 것으로 나타났다(Stewart-Brown, 2021^[16]). 의료 서비스를 이용하는 사람을 대상으로 실시한 연구에 따르면 응답자는 정신건강 설문조사에서 일련의 부정적인 항목을 질문하는 것에 대해 ‘화가 난다’고 답했다. 특히 양호한 정신건강에 초점을 맞춘 WEMWBS 문항을 선호하는 것으로 나타났다(Crawford 외, 2011^[111]).

표 3.3. 유사한 문항에 대한 응답의 상관관계는 문구가 긍정적인지 또는 부정적인지에 따라 달라진다

서로 다른 정신건강 선별도구에서 유사한 단어로 구성된 문항 간의 상관관계

| 문항 문구 | | | |
|---------|--|---------------------------------------|--------------------------------------|
| | GHQ-12 | SF-12 | SWEMWBS |
| | 문제에 직면할 수 있었음 | | 문제에 잘 대처함 |
| | 불행하고 우울한 느낌 | 낙담하고 우울한 느낌 | |
| | | 차분하고 편안한 느낌 | 편안한 느낌 |
| 응답 상관관계 | | | |
| | 불행/우울한 느낌 (GHQ-12 및 SF-12): 0.67 | 문제에 직면 (GHQ-12 및 SWEMWBS): 0.36 | 편안한 느낌 (SF-12 및 SWEMWBS): 0.56 |

주: 상관관계는 동일한 중단적 조사에서 서로 다른 정신건강 선별 도구 또는 척도에 포함되어 있는 유사한 문구로 구성된 각 쌍별 문항의 피어슨 상관계수(Pearson correlation coefficient)를 보여준다. 특정 도구에 대한 자세한 내용은 부속서 2.B를 참조한다.

출처: 에식스대학교 사회경제연구소(2022^[70]), *Understanding Society: Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009* (데이터베이스). 15th Edition. UK Data Service. SN: 6614, <http://doi.org/10.5255/UKDA-SN-6614-16>, from wave 10 only (2018년 1월 - 2020년 5월)를 기반으로 OECD에서 계산함.

Box 3.10. 주요 메시지: 문항 구성

- OECD 연구에 따르면 가구 조사에서 주관식 문항의 순서는 응답에 영향을 미칠 수 있다. 따라서 설문조사 문항을 일관된 순서로 유지하는 것이 중요하며, 가능하면 해당 질문은 설문조사 초반부에 배치해야 한다.
- 영국의 사회 이해 설문조사(*Understanding Society survey*)에 따르면 어떤 개념이 부정적 또는 긍정적 시각으로 구성되었는지에 따라 응답이 달라질 수 있다.
- 정신건강 관련 의료 서비스를 이용하는 사람 중 일부는 부정적 감정보다는 긍정적 감정에 초점을 맞춘 설문조사 도구를 선호하는 것으로 나타났다.

설문조사 모드가 응답자의 답변에 영향을 미치는가?

설문조사 모드, 즉 응답자로부터 데이터를 수집하는 방식은 응답자가 문항을 처리하고 응답하는 방식과 응답자가 공개하는 정보의 양에 영향을 미칠 수 있다. 설문조사 모드에 따라 응답이 달라지는 주요 원인 중 하나는 사회적 바람직성 편향으로, 응답자가 타인에게 호의적인 인상을 주거나 일반적인 사회적 규범에

부합하는 응답을 제공하려는 경향에 기인한다. 사회적 바람직성은 ‘인상 관리’와 ‘자기 기만’이라는 두 가지 요소가 있다(Paulhus, 1984_[112]). 연구에 따르면 응답자는 금기시되는 주제는 과소 보고하고 사회적으로 바람직한 행동을 과대 보고하는 것으로 나타났다(Krumpal, 2013_[59], Presser 및 Stinson, 1998_[113]). 사회적 바람직성 편향은 면담자가 데이터를 수집하는 방식, 즉 직접 대면하거나 전화 또는 인터넷을 통해 데이터를 수집하는 방식에 따라 다양한 방식으로 나타날 수 있다.

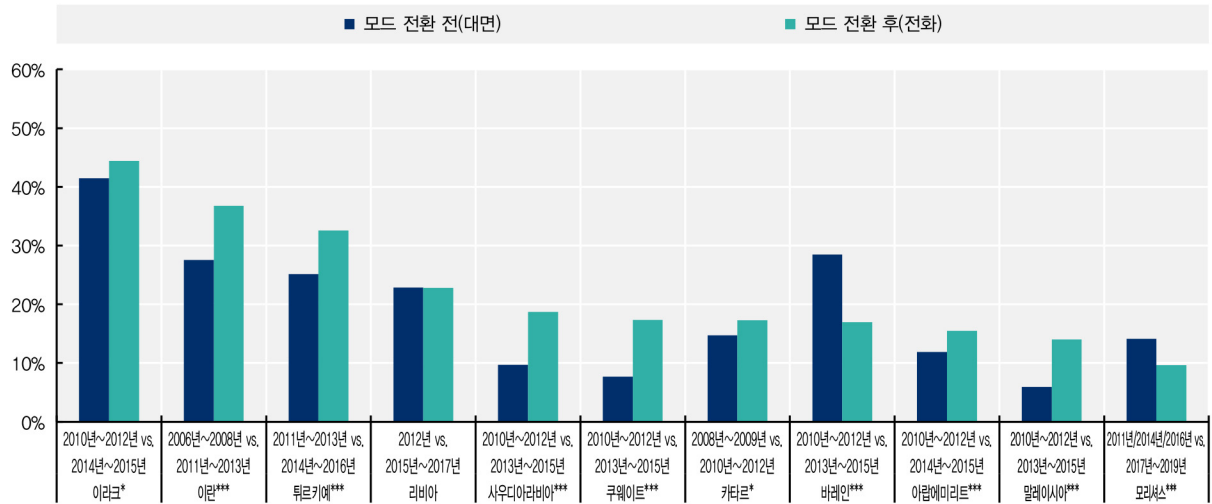
연구에 따르면 응답자는 면담자가 실시하는 설문조사에서 자기기입식 설문조사에 비해 더 우수한 신체적, 정신적 건강 결과를 보고할 가능성이 높은 것으로 나타났다. 또한 연구에 따르면 자기기입식 설문조사의 응답자는 불안 및 기분장애를 포함하여 낙인 찍힌 의학적 상태를 보고할 가능성이 낮은 것으로 나타났다(Krumpal, 2013_[59], Latkin 외, 2017_[114]). 노르웨이의 한 연구에 따르면 응답자가 면담자(대면 또는 전화)가 실시하는 설문조사에 비해 자기기입식 설문조사에서 불안과 우울 증상을 보고할 가능성이 더 높았다(Moum, 1998_[115]). 특히 사회적 바람직성 편향은 교육 수준이 높은 청년층의 응답자에서 두드러지게 나타났다. 응답자가 컴퓨터에 직접 입력하는 조사 방법(Computer assisted self-interviewing, CASI)과 종이와 연필을 이용한 조사 방법(Interviewer-administered paper-and-pencil, I-PAPI)을 비교한 또 다른 연구에 따르면 응답자는 면접관이 실시하는 설문조사보다 자기기입식 설문조사에서 세계보건기구의 복합 국제진단 면담 도구(WHO-CIDI)로 측정된 정신건강 증상을 보고할 가능성이 더 높았다(Epstein, Barker 및 Kroutil, 2001_[116]). 이는 주요우울상화와 범불안장애의 경우에서도 큰 차이를 보였다¹³.

면담자가 실시하는 설문조사의 경우 응답자가 전화나 인터넷을 통한 조사에서 면담자와 대면하여 실시하는 조사에 비해 높은 수준의 정신건강 결과를 보고하는지에 대한 근거가 엇갈리고 있다. 캐나다 지역 건강조사(CCHS)에 따르면 일부 신체건강 지표는 설문조사 모드의 영향을 받지만, 정신건강 결과는 영향을 받지 않은 것으로 나타났다¹⁴ (St-Pierre 및 Béland, 2004_[117]). 미국의 한 연구에서는 응답자가 면담자와 대면하여 실시하는 설문조사보다 전화 면담을 통한 설문조사에서 사회적 바람직한 행동이 더욱 두드러지는 정반대의 결과를 발견했다(Holbrook, Green 및 Krosnick, 2003_[118]). 마지막으로 면담자 감독하에 이루어지는 웹 기반 평가와 종이 및 연필을 이용한 조사 방법을 비교한 최근 메타 분석에 따르면 웹 기반 설문조사의 자기보고 설문지에서 사회적 바람직성과 관련하여 이점을 제공하지 않으며, 조사 모드가 정신건강 증상을 보고하는데 영향을 미치지 않는 것으로 나타났다(Gnambs 및 Kaspar, 2017_[119]).

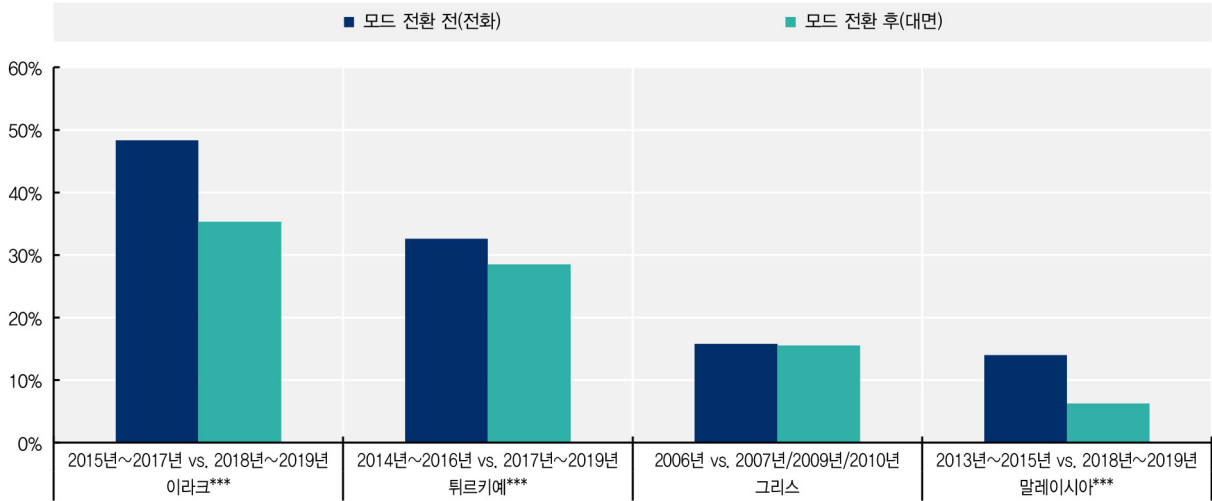
면담자가 실시하는 설문조사 모드가 응답자의 답변에 미치는 영향은 그림 3.7에 설명되어 있다. 이 그림은 Gallup World Poll 데이터에서 부정적 정서 균형, 즉 전날 긍정적인 감정보다 부정적인 감정을 더 자주 경험했다고 보고한 인구의 비율을 나타낸다. Gallup은 OECD 38개국을 포함하여 150개 이상의 국가에서 매년 설문조사를 실시하고 있다. 많은 OECD 국가에서는 전화 설문조사를 통해 데이터를 수집하지만 라틴 아메리카, 중동, 아시아, 아프리카 및 구소련 등 많은 국가에서는 대면 설문조사가 일반적이다¹⁵. 소수의 국가에서 지난 10년 동안 조사 방식이 변경되어 대면 조사에서 전화 조사로 또는 그 반대로 전환되었다(실제로 이라크, 튀르키예 및 말레이시아 등 일부 국가의 경우 여러 번 전환함). 그림 3.7은 대면 조사에서 전화 조사로 전환한 후 11개 국가 중 8개국에서 부정적 정서 균형이 증가했음을 보여준다(패널 A). 마찬가지로 전화 조사에서 대면 조사로 전환한 후 4개국 모두에서 부정적인 정서 균형이 개선되었다(패널 B). 이는 응답자가 면담자와 대화를 통해 대화할 때 사회적 바람직성 편향의 영향을 더 많이 받아 전화 조사에서 웰빙을 과대 보고할 수 있다는 연구 결과와 일치한다(Holbrook, Green 및 Krosnick, 2003_[118])¹⁶.

그림 3.7. 대면 조사에서 전화 조사로의 전환은 부정적 정서 균형의 악화와 관련된다

패널 A: 다양한 연도별 부정적 정서 균형 인구 비율,
설문조사 모드 전환(대면 조사에서 전화 조사로 전환) 전후 3개년 통합 평균, 11개국



패널 B: 다양한 연도별 부정적 정서 균형 인구 비율,
설문조사 모드 전환(전화 조사에서 대면 조사로 전환) 전후 3개년 통합 평균, 4개국



주: ***는 설문조사 모드 변경 후 통계적으로 유의미한(5% 수준) 결과 변화를 경험한 국가를 나타내며, *는 10% 수준에서 통계적으로 유의미한 변화를 경험한 국가를 나타낸다. 3개년 평균(설문조사 모드 변경 전 3년과 변경 후 3년)이 표시되며 모드 변경 전후로 데이터 수집 기간이 충분하지 않은 국가는 예외로 한다. 해당 경우에는 1개년 또는 2개년 평균이 표시된다. 이라크와 아랍에미리트에서는 설문조사 모드가 변경된 연도인 2013년에 부정적 정서 데이터를 수집하지 않았다.
출처: Gallup World Poll(n.d.[120]) (데이터베이스), <https://www.gallup.com/178667/gallup-world-poll-work.aspx>를 기반으로 OECD에서 계산함.

StatLink <https://stat.link/ngkd9r>

상자 3.11. 주요 메시지: 설문조사 모드 효과

- 응답자는 면담자가 실시하는 설문조사에 비해 자기기입식 설문조사에서 정신건강을 더 열악한 수준으로 보고할 가능성이 높다.
- 면담자가 실시하는 설문조사의 경우 정신건강 결과가 모드에 영향을 받는지에 대해 상반된 근거가 있다.
- 설문조사 모드의 일관성을 유지하는 것이 권장되며 데이터 수집 활동 사이에 설문조사 모드가 변경되는 경우 이를 명시적으로 표시해야 한다.

정신건강 모듈 사용 시 데이터 수집기관에서 적용해야 하는 추가 프로토콜 또는 절차는 무엇인가?

면담자 교육은 모든 설문조사에서 응답의 질을 높이는 데 매우 중요하다. 그러나 정신건강 측정은 주제의 민감한 특성으로 인해 추가적인 문제가 발생한다. 면담자 교육을 받은 면담자는 일반적으로 높은 응답률과 응답의 질 향상에 기여하지만, 면담자는 해당 정보를 수집하는 것이 왜 중요한지, 수집된 정보가 어떻게 사용되는지에 대해 응답자에게 적절하게 설명할 수 없는 경우 문항에 대한 응답을 수집하는 데 어려움을 겪을 수 있다. 경우에 따라 응답자는 공공 기관에서 이러한 유형의 정보를 수집하려는 이유를 이해하지 못할 수도 있다. 정신건강 관련 문항에 대한 응답자의 태도와 관련 위험을 관리하려면 면담자가 해당 문항을 통해 측정하려는 개념뿐만 아니라 수집된 정보가 어떻게 사용될 것인지에 대해 충분히 숙지하고 있는 것이 필수적이다. 이는 면담자가 응답자와 강한 유대감을 형성하여 응답률과 응답의 질을 개선하는 데 도움이 될 수 있다.

응답자와 면담자의 관계도 중요하다. 임상 환경에서 정신건강 관련 의료 서비스를 이용하는 집단을 대상으로 실시한 연구에서 응답자는 설문지는 ‘설문지를 사용하는 의사만큼만 좋은 것’이라고 강조했다(Crawford 외, 2011_[111]). 실제로 응답자는 설문조사의 내용이나 길이보다 면담자가 가장 중요하다고 답했다. 미국의 한 연구에 따르면 응답자는 전화 조사보다 대면 조사에서 불법 약물 사용에 대한 민감한 정보를 공개할 가능성이 더 높았으며 그 차이는 백인계 미국인에 비해 흑인계 미국인에서 더 두드러지게 나타났다(Aquilino, 1994_[121]). 노르웨이의 한 연구에서는 면담자의 성별과 연령이 정신건강 증상 보고에 미치는 영향은 미미한 것으로 나타났지만, 젊은 연령대의 남성 면담자는 다른 연령대의 면담자나 여성 면담자에 비해 증상에 대한 보고가 더 적은 것으로 나타났다(Moum, 1998_[115]). 이는 역으로도 작용하여 면담자와 응답자의 유대감이 강할수록 응답자는 정보를 더 많이 공개할 수 있다¹⁷.

최근 연구에서는 정신의학적 장애 또는 정신건강 상태를 실제 경험한 연구자의 파이프라인을 구축하여 설문조사 설계 및 데이터 수집 과정에 참여시켜야 할 필요성에 대해 조명했다(Jones 외, 2021_[122], Banfield 외, 2018_[123], Hancock 외, 2012_[124]). 피면담자와 동일한 공동체에서 조사원을 모집하는 동료 면담 기법은 면담자와 피면담자 간의 신뢰 수준을 높이고 접근이 어려운 인구집단에 대한 고품질 데이터를 수집하는 데 효과적인 방법임을 뒷받침하는 강력한 근거가 있다(Dewa 외, 2021_[125], Warr, Mann 및 Tacticos, 2011_[126], Hancock 외, 2012_[124]). 또한 정신건강 맥락에서 실제 정신질환 경험이 있는 사람을 연구에 참여 시키면 연구 결과의 신뢰도와 정책 채택 가능성을 모두 높일 수 있다(Scholz 외, 2021_[127]).

자살 또는 자살 사고에 대한 문항은 정신건강 위험군에 필요한 지원을 제공하기 위해 신중하고 적절하게 설계된 절차가 필요하다(Lakeman 및 FitzGerald, 2009_[128]). PHQ-9의 마지막 항목은 자살 생각과 자살 사고에 대해 질문하는 문항으로 인구 조사에서 제외되는 경우가 많다. 바로 이러한 이유로 유럽건강면접조사(EHIS)와 미국 국민건강면접조사(NHIS)에서는 PHQ-8이 대신 사용된다.

자살에 대한 문항을 포함하는 국가에서는 추가 프로토콜을 사용하는 경우가 많다. 예를 들어 호주의 비자살적 자해(non-suicidal self-injury, NSSI) 유병률에 관한 연구에서는 설문조사 문항의 민감한 특성을 고려하여 면담 대상 가구와 사전에 대량의 정보를 공유함으로써 윤리적으로 고려해야 할 사항을 축소하고 무응답률을 낮추는 데 도움이 되었다(Taylor 외, 2011_[129]). 캐나다 정부는 정신건강 및 치료 접근성 조사(Mental Health and Access to Care Survey, MHACS)를 실시하면서 응답자와 면담자 모두에게 정신건강 자원을 제공하고 조사원에게는 설문조사 수행으로 인해 유발될 수 있는 정신건강에 악영향이나 스트레스를 완화할 수 있도록 직원 지원 서비스도 제공한다(OECD 설문지에 대한 응답, 2022).

상자 3.12. 주요 메시지: 면접관 교육

- 응답자는 면담자가 편안하다고 느낄 경우 면담에 참여하고 질문에 진실하게 답변할 가능성이 높다. 조사원 교육은 응답자와의 친밀감 및 신뢰 구축에 중점을 두어야 한다.
- 설문조사에 자살 및 자살 충동에 대한 문항을 포함하려면 신중한 절차 및 지원 관행을 마련해야 한다. 가구 조사의 경우 윤리적인 이유로 해당 유형의 문항은 포함시키지 않도록 권장된다.

분석

정신건강 데이터는 가능한 한 정확하고 일관성 있게 수집되어야만 정책 입안자에게 유용한 데이터 분석이 가능하다. 본 보고서에 설명된 많은 데이터는 이분법이 아니며, 즉 결과 변수가 척도에서 측정된다. 이는 선별 도구와 복합 척도에도 해당되며 이러한 척도에는 여러 문항이 포함되어 있다. 그에 따라 최악의 정신건강에서 최상의 정신건강까지 코딩 및 점수가 매겨진다. 일반적인 정신건강 상태 도구는 일반적으로 단일 문항이지만 응답 옵션은 이분법이 아니며, 일반적으로 리커트 척도를 사용한다(응답 옵션의 정확한 수와 문구는 척도마다 다름, 표 2.11 참조).

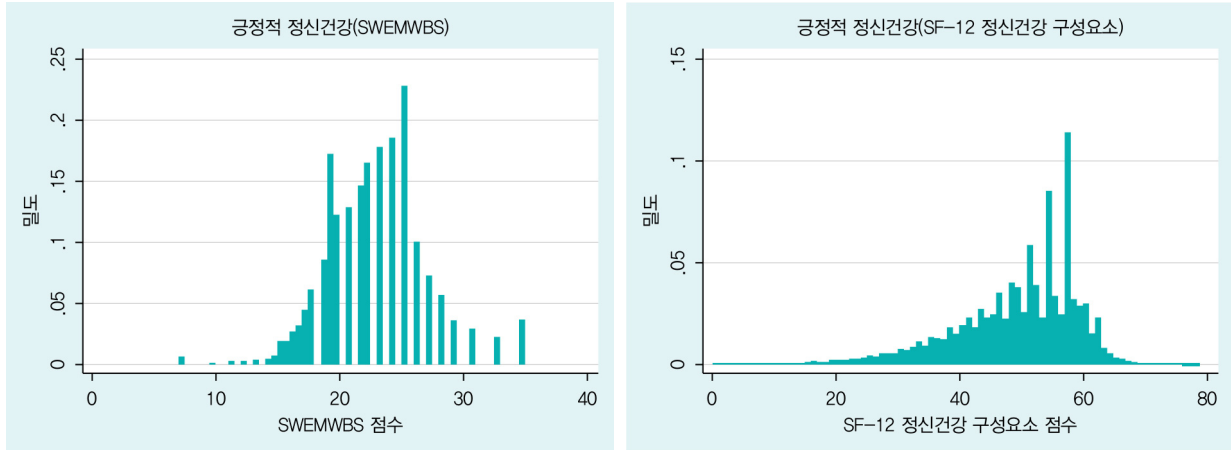
준거점수를 사용하는 경우와 연속 측정을 사용하는 경우의 장단점은 무엇인가?

정신건강 결과에 대한 연속 측정은 인구 정신건강에 대해 보다 상세하고 세밀한 정보를 제공하는 등 많은 이점이 있다. 또한 응답 분포가 바닥효과 또는 천장효과 없이 정상적으로 분포된 경우¹⁸ 결과에 대한 모수적 분석이 가능하여 특정 정책이나 정책적 개입의 영향에 대해 더욱 정확한 분석이 가능해진다. 예를 들어 긍정적 정신건강을 위한 선별 도구에 대한 연구에 따르면 (S)WEMWBS에서 수집된 데이터는 특정 정신질환에 초점을 맞춘 선별 도구보다 정규 분포에 더 가까운 분포를 보인다(Shah 외, 2021_[17])¹⁹. 실제로 이는 제10차 UKHLS 조사에 포함된 세 가지 정신건강 도구의 밀도 그래프를 그림 3.8에서 시각적으로 확인할 수 있다. SWEMWBS, SF-12(MHC-12)의 정신건강 구성요소, GHQ-12 등 세 가지 정신건강 도구 중에서 SWEMWBS는 정규 분포에 가장 가까우며 그 다음으로 SF-12가 정규 분포에 가깝고 GHQ-12는 유의미한 바닥효과를 보여준다. MHI-5에 대한 별도 연구 결과는 긍정적으로 편향되어 있는 것으로 나타났고, 이는 긍정적 정신건강 수준이 높은 사람보다 정신건강이 열악한 사람을 더 효과적으로 구별할 수 있음을 의미한다(Elovanio 외, 2020_[6], Thorsen 외, 2013_[130]).

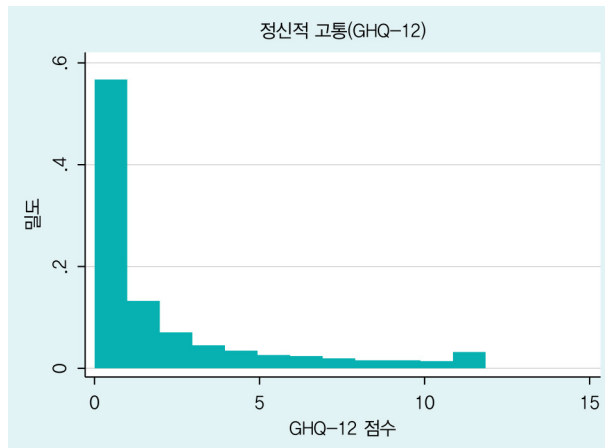
그림 3.8. 긍정적 정신건강 척도는 정규 분포에 더 가까울 수 있다

패널 A: 긍정적 정신건강에 대한 가중 밀도 그래프 (SWEMWBS), 영국, 2012년~2020년

패널 B: 긍정적 정신건강에 대한 가중 밀도 그래프(SF-12 정신건강 구성요소), 영국, 2010년~2020년



패널 C: 정신적 고통에 대한 가중 밀도 그래프(GHQ-12), 영국, 2010년~2020년



주: 가중 점수를 나타내는 밀도 그래프로 패널 A의 경우 SWEMWBS[9.5점(정신적 웰빙 악화)에서 35점(정신적 웰빙 개선)까지], 패널 B의 경우 SF-12의 정신건강 구성요소[0점(기능 낮음)에서 100점(기능 높음)까지], 패널 C의 경우 GHQ-12[0점(정신건강 개선)에서 12점(정신건강 악화)까지]가 가장 낮은 점수이다. SF-12 및 GHQ-12의 데이터는 제2차~제10차 *사회의 이해(Understanding Society)*에서, SWEMWBS 데이터는 제4차, 제7차, 제10차 *사회의 이해(Understanding Society)*에서 추출되었다. 개별 선별 도구에 대한 자세한 내용은 부속서 2.B를 참조한다.

출처: 에식스대학교 사회경제연구소(2022^[70]), *Understanding Society: Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009* (데이터베이스). 15th Edition. UK Data Service. SN: 6614, <http://doi.org/10.5255/UKDA-SN-6614-16>, from wave 10 only (2018년 1월 - 2020년 5월)를 기반으로 OECD에서 계산함.

정규 분포는 회귀 분석에 유용하지만 정신건강 정보를 미시적 수준(예: 환자가 위험군이며 더 많은 지원이 필요한지 확인하기 위해 선별 면담을 수행하는 주치의)이나 거시적 수준(예: 시간 경과에 따른 위험 변화를 추적해야 하는 정부기관)에서 유용하게 사용하려면 준거점수를 사용하여 결과를 범주로 그룹화하는 것이 유용한 경우가 많다. 이러한 범주는 사용된 선별 도구 및 채점 방식에 따라 다르지만 일반적으로 ‘우울증 위험군’, ‘불안 위험군’, ‘주요우울장애’, ‘심각한 심리적 고통’, ‘심리적 변영’ 등을 포함한다. 이러한 범주는 정신건강이 웰빙의 다른 측면과 어떻게 상호작용하는지(예: 불안을 경험하는 취업자 또는 실업자 비율, 우울증 비위험군 대비 우울증 위험군의 사회적 관계의 질 등) 이해하는 분석에도 유용할 수 있다(OECD, 2021^[57]).

일반적으로 임계값이 임의적일 수 있다는 점에서 임계값을 사용하는 것에 대해 비판적인 시각이 있다. 그러나 정신건강 선별 검사의 경우 엄격한 검증 과정을 통해 임계값이 설정되며, 수신자조작특성(ROC) 분석을 통해 측정의 민감도와 특이도를 모두 극대화하는 준거점수를 결정한다(상자 3.3 참조)²⁰. 또한 준거점수는 일련의 선별도구 문항에 대한 응답을 심층진단면담의 결과와 유사한 것, 즉 특정 정신건강 상태에 대해 위험한 것으로 정의한다는 점에서도 유용하다. PHQ, GAD, 케슬러 및 CES-D 조사에는 표준 준거점수가 검증되어 있다(Kessler 외, 2002^[89], Kroenke 외, 2007^[44], Moriarty, Zack 및 Kobau, 2003^[21], Manea, Gilbody 및 McMillan, 2015^[131], Kroenke 외, 2009^[12], Spitzer 외, 2006^[11]). 임상 표본에서 ‘우울증 사례’를 나타내는 기존 CES-D 준거점수는 16점이지만 해당 임계값은 비임상 표본에서 높은 위양성률(false positive)을 생성하는 것으로 알려져 있다(Eaton, 2004^[132], Santor 및 Coyne, 1997^[133]). GAD-7도 준거점수가 정해져 있으나 연구에 따르면 범불안장애 위험군 비율을 식별하는 데 성능이 더 우수하며 사회불안장애와 같은 다른 유형의 불안장애를 포착하는 데 있어 성능이 저하되는 것으로 나타났다(Beard 및 Björgvinsson, 2014^[134], Sunderland 외, 2019^[33]).

GHQ-12는 일반적인 정신건강 상태를 선별하는 데 일반적으로 사용되지만 높은 수준의 위양성을 생성하는 것으로 밝혀졌다. 한 연구에 따르면 정신장애가 있는 것으로 확인된 사람의 절반 가량이 위양성(양성예측도 0.53)인 것으로 나타났다(Schmitz, Kruse 및 Tress, 2001^[4]). MHI-5 또는 GHQ-12와 같은 다른 정신건강 선별 도구는 일반적인 정신장애 사례를 정의하기 위해 검증된 표준 준거점수가 개발되지 않았다. 이러한 척도에는 국제적으로 비교 가능한 준거점수가 없을 수 있지만 여러 연구에서 검증되었다. 예를 들어 Berwick 외(1991^[38])에서 임상 면담을 표준 기준으로 사용하여 우울증 척도로서 MHI-5를 검증한 결과, 최적의 준거점수가 52점으로 나타났다²¹. 후속 연구에서는 MHI-5가 우울증 및 일반적인 기분장애에 대한 선별 도구로 성능이 우수하나 불안, 신체형 장애 및 약물 사용장애의 척도로서는 성능이 저하된다는 사실이 입증되었다(Rumpf 외, 2001^[135], Strand 외, 2003^[7], Thorsen 외, 2013^[130]).

일부 도구는 진단 목적에 따라 허용되는 준거점수가 여러 개이므로 채점 방식에 따라 유병률 추정치가 달라질 수 있다. 그림 3.9는 OECD 국가 중 유럽 22개국의 PHQ-8 점수에 대한 밀도 그래프로 0점(우울증 위험 가장 낮음)에서 25점(우울증 위험 가장 높음)까지의 범위를 가진다. 수직선에는 서로 다른 검증된 임계값이 표시된다. 10점 이상은 주요우울장애의 위험을 나타낸다(검은색으로 표시)(Kroenke 외, 2008^[136]). 다른 임계값 범주에서는 5점~9점을 경증 우울증 위험, 10점~14점을 중등도 우울증 위험, 15점~19점 중증 우울증 위험으로, 20점 이상을 최중증 우울증 위험으로 간주한다(Kroenke, Spitzer 및 Williams, 2001^[137]). Eurostat에서 사용하는 또 다른 채점 방식(그림 3.9에 표시되지 않음)은 원점수만을 기준으로 하지 않고 개별 문항에 대한 응답자의 답변을 통해 주요 우울 증상을 정의한다²². 세 가지 측정 방식은 다음과 같이 동일한 데이터 세트에서 서로 다른 유병률 추정치를 도출한다. (1) 주요우울장애 위험군 6.9%, (2) 경증 우울증 위험군 15.2%, 중등도 우울증 위험군 2.9%, 중증 우울증 위험군 1.7%, 최중증 우울증 위험군 0.8%, (3) 주요 우울 증상이 있는 경우 3.1%²³.

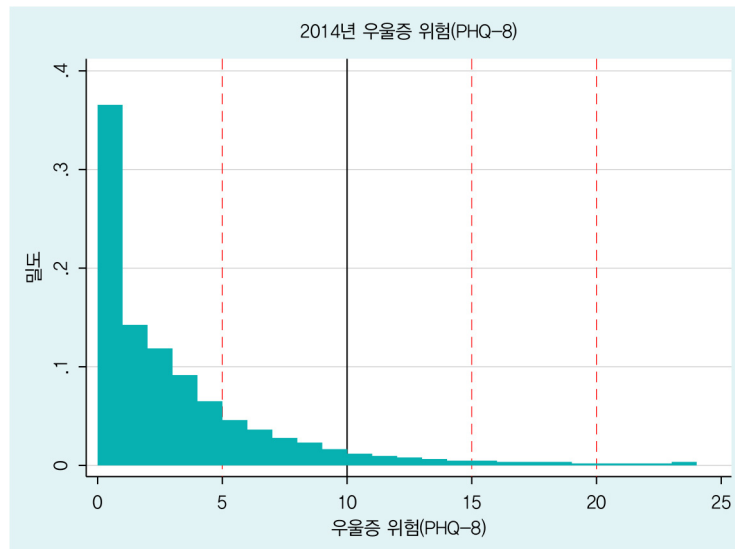
심리적 웰빙에 대한 임상적 표준 기준은 없지만, 긍정적 정신건강 복합 척도의 경우 사용자의 요청에 따라 준거점수를 개발하기도 했다. (S)WEMWBS에 대한 두 가지 주요 접근 방식이 제시되었는데, 하나는 통계적 접근 방식이고 다른 하나는 벤치마킹 방식이다²⁴. 첫 번째 접근 방식에서 +/- 1 표준 편차에서 준거점수가 권장되며, 이를 통해 인구의 약 15%가 높은 수준의 웰빙에, 또 다른 15%가 낮은 수준의 웰빙에 포함된다. 두 번째 접근 방식에서는 우울증과 불안의 증상을 포착하는 척도를 기준으로 (S)WEMWBS의 준거점수를 벤치마킹한다(기분장애의 선별 도구로 사용되는 긍정적 정신건강 도구에 대한 자세한 설명은 아래 참조). 여러 연구에서 WEMWBS를 CES-D, PHQ-9 및 GAD-7에 벤치마킹하여 임상적 우울증 가능성, 발생 가능한 우울증 또는 경증 우울증(또는 불안)의 위험을 나타내는 WEMWBS 척도의 준거점수를 제시했다.

모든 측면을 종합해 볼 때 정신적 웰빙 수준이 높은 상위 15%의 사람을 식별하는 데 WEMWBS 척도에서는 60점, SWEMWBS 척도에서는 28점의 준거점수를 사용할 수 있다고 권장된다. 그러나 높은 수준의 정신적 웰빙에 대한 임상적 측정이 존재하지 않아 이러한 임계값은 임의적이라는 점에 주의해야 한다(Warwick Medical School, 2021^[78]).

MHC-SF는 세 가지 하위 척도(정서적, 사회적, 심리적 웰빙)로 구성되며, 이를 점수화하여 개인을 번영, 쇠약함, 중간 수준(앞의 두 범주에 속하지 않는 경우) 등 세 가지 범주로 분류할 수 있다(Lamers 외, 2011^[20]). 그러나 캐나다와 덴마크의 설문조사에 따르면 대다수의 인구가 번영으로 분류될 만큼 높은 점수를 받았으며, 이는 번영과 쇠약함이 소수의 인구를 대표하고 평균에서 벗어난다는 이론에 반하는 결과이다. 이러한 점은 측정의 민감도를 개선하기 위해 보다 보수적인 채점 기준이 필요할 수 있음을 시사한다(Santini 외, 2020^[48]).

그림 3.9. 채점 방식이 다르면 유병률 추정치가 달라질 수 있다

우울증 위험군 분포를 보여주는 밀도 그래프(PHQ-8), OECD 국가 중 유럽 22개국, 2014년



주: OECD 국가 중 유럽 22개국의 PHQ-8 점수에 대한 가중 밀도 그래프로 점수 범위는 0점(우울증 위험 가장 낮음)에서 25점(우울증 위험 가장 높음)까지이다. 수직선은 문헌에서 설정한 검증된 준거점수를 나타낸다. 검은색 굵은 실선으로 표시된 수직선은 점수가 10점 이상이면 주요우울장애의 위험을 나타낸다(Kroenke 외, 2008^[136]). 빨간색 점선으로 표시된 수직선은 점수가 5점~9점은 경증 우울증 위험, 10점~14점은 중등도 우울증 위험, 15점~19점은 중증 우울증 위험, 20점 이상은 최중증 우울증 위험을 나타낸다(Kroenke, Spitzer 및 Williams, 2001^[137]).

출처: 제2차 유럽건강면접조사(EHIS)(n.d.^[66]) (데이터베이스), [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:European_health_interview_survey_\(EHIS\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:European_health_interview_survey_(EHIS)). 데이터를 기반으로 OECD에서 계산함.

상자 3.13. 주요 메시지: 연속 측정 및 준거점수

- 이분법적인 결과 변수가 아닌 연속적인 결과 변수를 제공하는 정신건강 도구를 사용하면 인구 수준의 정신건강을 보다 세밀하게 파악할 수 있다.
- 긍정적 정신건강 복합 척도가 심리적 고통을 측정하는 척도보다 정규 분포에 가깝다는 근거가 있다.
- 준거점수는 연구자와 정책 입안자에게 정신건강 위험군과 비위험군을 명확하게 구분할 수 있는 범주를 제공한다. 정신건강 문제 선별 도구의 준거점수는 민감도와 특이도를 극대화하기 위해 임상 진단에 대해 검증되었지만 심리적 안녕감의 경우 이러한 표준 기준이 마련되어 있지 않다.

- 민감도 및 특이도 극대화를 위한 검증에도 불구하고 준거점수는 위양성을 제공하거나 일부 인구집단에 적합하지 않을 수 있다.
- 동일한 선별 도구에 대해 다른 채점 방식을 적용하면 유병률 추정치가 달라질 수 있으므로 일관성을 유지하기 위해 주의를 기울여야 한다.

결론

본 장에서는 인구 정신건강을 측정하는 데 가장 적합한 도구를 생각할 경우, 중요하게 고려해야 하는 사항에 대해 다룬다. 설문조사 데이터와 마찬가지로 여러 고려사항이 존재하며, 시간 경과에 따른 변화와 집단 간 편차를 해석할 시에는 주의가 필요하다. 정신건강에 대한 낙인과 차별적인 시각은 보고된 데이터의 편향성을 유발할 수 있다. 또한 데이터의 품질 및 정책 관련성을 보장하기 위해 실제 설문조사 설계 경험자의 관점을 통합하는 것이 중요하다. 그러나 본 장에서 검토한 근거에 따르면 기존 정신건강 도구가 유용하고 정책과 관련된 결과를 제공하는 것으로 나타났다. 정신건강 위기의 시급성과 각국 정부의 대응 우선순위를 고려하여 기존의 모범 사례에 따라 고품질 정신건강 데이터를 수집하는 것은 매우 중요하다. 측정 관련 미해결된 사안에 대한 연구는 인구 정신건강에 대한 지속적인 모니터링과 함께 진행될 수 있다.

현재 모든 OECD 국가는 어떤 형태로든 인구 정신건강을 측정하고 있으며 이미 집단 간 종단적 비교 연구를 진행하고 있다. 일부 도구의 경우 시간 경과에 따른 변화에 대한 민감도를 검증하기 위해 추가적인 연구가 필요하다. 그러나 코로나19 팬데믹 기간에 높은 빈도로 수집된 데이터는 많은 정신건강 측정이 실제로 변화에 민감하다는 사실을 입증했다. 코로나19 이전에는 범불안장애 및 우울장애와 같은 일반적인 정신질환의 비율이 최근 몇 년간 안정적으로 유지되어 왔다는 점을 강조하는 정책적 논의가 있었다. 그러나 현재는 팬데믹 발발 후 처음 2년 동안 심리적 고통의 비율이 급격히 증가했으며, 이러한 결과가 OECD 국가에서 수집된 데이터에서 포착되었다는 점에 광범위한 공감대가 형성되어 있다.

앞으로의 과제는 데이터 수집을 통일화시키고 이미 진행 중인 계획의 품질 개선을 위한 권장사항을 제공하는 것이다. 제2장에서 소개된 2022년 OECD 설문조사의 결과는 불안 증상을 측정하는 통일된 방식의 부재, 정서 및 행복 측정 도구의 표준화 부재, 우울증이나 불안이 아닌 특정 정신건강 상태를 측정하는 도구의 불균일한 사용 등 국가별로 수집된 정신건강 결과 유형에 여전히 격차가 있음을 보여준다. 이러한 영역에 대한 권장사항은 아래에 제시되어 있으며, 데이터 수집의 실용성을 고려하여 새로운 설문조사 항목은 짧게 작성해야 한다는 점을 강조한다.

다양한 도구의 통계적 품질, 응답 부담 및 비용(문항 길이로 대신함)에 대한 비교 평가와 기존 데이터 수집 관행에 대한 정보(표 3.1)를 기반으로 **국가 통계청이 가구, 사회 및 건강 조사에 채택할 특정 정신건강 결과 측정 도구를 포함할 것을 권장한다.** 이러한 권장사항은 특히 이전 진단 및 경험한 증상이나 2013년 *OECD 주관적 웰빙 측정 지침*의 측정 항목(예: 삶의 만족도)과 관련하여 인구 정신건강 결과를 파악하기 위해 이미 사용하고 있는 다른 도구를 단계적으로 폐지하는 것을 의미하지 않는다. 오히려 해당 지표들은 보다 국제적으로 통일화된 인구 정신건강 결과 지표를 구축할 수 있도록 도움을 줄 수 있다.

- **정신건강 문제 우선순위 권장사항:** 환자 건강 설문지-4(PHQ-4)는 정신건강 설문조사에서 정기적으로 수집하는 PHQ-8 또는 PHQ-9와 함께 수행의 빈도수가 높은 설문조사에 포함될 수 있다. PHQ-4는 확장형 척도인 PHQ-9에서 추출한 우울증 문항 2개와 GAD-7 선별 도구에서 추출한 불안 문항 2개를 결합한 것이다. 해당 척도는 가장 일반적인 정신건강 상태인 우울증과 불안 중 어느 하나에만 초점을 맞추지 않고 두 가지를 모두 다루고 있다. 또한 4개의 문항으로 구성되어 있어 모듈이 비교적 짧고 응답 부담이 적다. OECD 국가의 81%가 이미 PHQ-8 또는 PHQ-9를 실시하고 있으며, 이는 PHQ-4의 우울증 문항과 연계할 수 있는 추세 데이터가 있음을 의미한다²⁵. PHQ-8 또는 PHQ-9 및 GAD-7은 특정 정신건강

조사에서 유지될 수 있으며 PHQ-4는 길이가 짧기 때문에 일반적으로 빈번하게 수행되는 조사에 채택할 수 있다.

- **긍정적인 정신건강-권장사항:** 국가별로 표준화된 방식으로 긍정적 정신건강의 정서적 및 자아실현적 측면을 측정하는 데 **WHO-5** 또는 **SWEMWBS**를 사용할 수 있다. 이러한 권장사항은 주로 국가별 측정 관행의 추세를 기반으로 한다. **WHO-5**는 긍정적 정서를 측정하는 도구로 비교적 짧고 쉽게 시행할 수 있으며, OECD 주관적 웰빙 측정 지침에 경험적 정서 모듈로 포함되어 있다(OECD, 2013^[11]). 여러 언어로 번역되어 제공되며 신뢰도와 타당도가 입증되었다. 현재 OECD 국가의 16%만이 사용하고 있지만 기업의 비재무적 성과를 측정하기 위한 개념적 프레임워크의 일부로 사용하거나(Siegerink, Shinwell 및 Žarnic, 2022^[138]) 환자에게 중요한 결과를 중심으로 의료 서비스를 제공하기 위해 환자보고지표조사(PaRIS)를 사용하는 등 다른 OECD 프로젝트에서 사용하도록 권장되고 있다(de Bienassis 외, 2021^[139]). **SWEMWBS**는 긍정적 정신건강의 정서적, 행복 및 사회적 관계 측면을 다룬다는 점에서 보다 포괄적인 도구에 해당한다. **SWEMWBS**는 **WHO-5**에 비해 문항이 2개 더 많다. **SWEMWBS** 또는 14개 문항으로 구성된 **WEMWBS**는 OECD 국가의 19%에서 채택하여 사용하고 있다. 이미 주관적 웰빙 또는 긍정적 정신건강 측정에 적극적으로 참여하고 있는 국가의 경우 **SWEMWBS** 내의 일부 지표 항목은 삶의 평가 및 사회적 관계의 양과 질 등의 개념을 측정하기 위한 기존 데이터 수집 노력과 중복될 수 있다[기존 OECD 권장사항 및 사례는 OECD(2020^[140]) 및 OECD(2013^[11]) 참조]. 이러한 경우에는 **WHO-5**가 정서만을 다룬다는 점에서 더 적합할 수 있다. 정서 및 행복을 구체적으로 측정하는 주제는 향후 주관적 웰빙에 관한 OECD 연구에서 계속 탐구될 것이다.
- **일반적인 정신건강 상태-권장사항:** 응답자의 **일반적인 정신건강 상태**에 대한 단일 문항은 한 국가의 전체 데이터 인프라 시스템에 걸쳐 다양한 설문조사에 포함될 수 있다. 일반적인 정신건강에 대한 단일 문항은 기존의 선별 도구에 비해 근거 기반이 부족하지만, 현존하는 연구 결과에 따르면 이는 유용하고 의미있는 척도임을 시사한다. 많은 OECD 국가에서 이미 자기보고 신체건강 데이터를 수집하고 있으므로 문항 구성 시 자기보고 **신체건강**과 자기보고 **정신건강**을 구별하는 것이 중요하다. 현재 일부 OECD 국가에서는 신체건강 및 정신건강을 모두 포착하는 자기보고 일반 건강 측정치를 수집하고 있으므로 이를 구분하여 수집할 것을 권장한다. 이를 국제적으로 비교 가능한 방식으로 수행하기 위해서는 추후 많은 연구와 조정을 통해 기존 국가별 관행에 맞추어야만 한다. 캐나다는 이러한 접근 방식을 일찍부터 채택했으며 자기보고 정신건강(SRMH) 문항 및 답변 옵션으로서 해당 프레임워크를 칠레와 독일에서 이미 채택했다. 또한, 이러한 유형의 문항에 대한 기존 근거의 상당 부분이 캐나다에서 도출되었다. 설문조사에 문항을 추가하고자 하는 다른 국가에서도 다음과 같은 프레임워크를 사용하는 것을 고려해볼 수 있다. “일반적으로 귀하의 정신건강은 어떻습니까? 훌륭하다/매우 좋다/좋다/보통이다/나쁘다.”

현재 약물 사용장애, PTSD, 강박장애, 섭식장애, 양극성장애 등 우울증과 불안 이외의 정신건강 상태에 대한 정보를 수집하는 도구를 사용하는 국가는 극소수이다. 이러한 결과는 구조화된 면담을 사용하는 모든 국가에서 다루어지지만(표 2.4 참조), 프랑스와 슬로베니아는 해당 질환의 진단 및 증상을 파악할 수 있는 도구가 포함된 세부 조사 모듈을 구현했다. 이러한 개념은 일반적인 정신질환의 일부가 아닌 별개의 질환으로 측정하는 것이 가치가 있으므로 향후 해당 분야의 권장사항 개발에 초점을 맞출 수 있는 통계적 협의가 이루어져야 한다.

일반적으로 동일한 유형의 도구를 여러 번 반복하여 시행하는 것보다 도구 유형과 정신건강 결과 측정을 다양화하는 것이 더 많은 정보를 얻고 제한된 자원을 보다 효율적으로 활용할 수 있다는 점에 유의해야 한다. 예를 들어 한 국가의 조사 인프라 전반에 걸쳐 우울증 또는 불안을 포착하기 위해 다양한 선별 도구를 구현하는 것보다 도구 영역 **내에서** 통일성을 이루는 것이 더 유용할 수 있다. 즉 우울증 또는 불안을 측정

하는 단일의 선별 도구를 선택한 다음 이를 단일 문항으로 구성된 도구로 보완하여 진단 여부, 증상 경험 등을 파악할 수 있다.

특히 본 보고서에서는 결과 측정을 설명하는 데 있어 정확성이 중요하다는 점을 강조했다. 각 도구는 특정한, 약간 차이점이 있는 인구의 정신건강 측면을 측정한다. 또한 개별 도구는 다양한 방식으로 점수를 매길 수 있으며 각 도구에 따라 정신건강 결과에 대한 추정치가 달라질 수 있다. 이러한 점은 정신건강의 의미와 측정 방법을 설명하는데 있어서 명확한 의사소통이 필요할 뿐만 아니라 통일성이 필요함을 암시한다. 이는 팬데믹 이후 정신건강이 최우선 국가적 의제로 부상한 가운데 더욱 중요한 사안이다.

참고문헌

- Ahmad, F. et al. (2014), “Single item measures of self-rated mental health: A scoping review”, *BMCI Health Services Research*, Vol. 14/398, pp. 1-11, <https://doi.org/10.1186/1472-6963-14-398>. [51]
- Ahn, J., Y. Kim and K. Choi (2019), “The psychometric properties and clinical utility of the Korean version of GAD-7 and GAD- ”, *Frontiers in Psychiatry*, Vol. 127/10, <https://doi.org/10.3389/fpsy.2019.00127>. [10]
- Ali, G., G. Ryan and M. De Silva (2016), “Validated screening tools for common mental disorders in low and middle income countries: A systematic review”, *PLoS ONE*, Vol. 11/6, <https://doi.org/10.1371/journal.pone.0156939>. [32]
- Aquilino, W. (1994), “Interview mode effects in surveys of drug and alcohol use: A field experiment”, *Public Opinion Quarterly*, Vol. 58/2, pp. 210-240, <https://doi.org/10.1086/269419>. [121]
- Banfield, M. et al. (2018), “Lived experience researchers partnering with consumers and carers to improve mental health research: Reflections from an Australian initiative”, *International Journal of Mental Health Nursing*, Vol. 27/4, pp. 1219-1229, <https://doi.org/10.1111/INM.12482>. [123]
- Beard, C. and T. Björgvinsson (2014), “Beyond generalized anxiety disorder: Psychometric properties of the GAD- in a heterogeneous psychiatric sample”, *Journal of Anxiety Disorders*, Vol. 28/6, pp. 547-552, <https://doi.org/10.1016/j.janxdis.2014.06.002>. [134]
- Berwick, D. et al. (1991), “Performance of a five-item mental health screening test”, *Medical Care*, Vol. 29/2, pp. 169-176, <https://doi.org/10.1097/00005650-199102000-00008>. [38]
- Bharadwaj, P., M. Pai and A. Suziedelyte (2017), “Mental health stigma”, *Economics Letters*, Vol. 159, pp. 57-60, <https://doi.org/10.1016/j.econlet.2017.06.028>. [73]
- Böhnke, J. and T. Croudace (2016), “Calibrating well-being, quality of life and common mental disorder items: Psychometric epidemiology in public mental health research”, *The British Journal of Psychiatry*, Vol. 209/2, pp. 162-168, <https://doi.org/10.1192/BJP.BP.115.165530>. [50]
- Botha, F., P. Butterworth and R. Wilkins (2021), “Taking the pulse of the nation: Validating a single-item measure of mental distress”, *Melbourne Institute Working Paper Series*, No. 6/21, Melbourne Institute of Applied Economic and Social Research, The University of Melbourne, <https://melbourneinstitute.unimelb.edu.au/publications/working-papers/search/result?paper=3821299>. [52]
- Bowe, A. (2017), “The cultural fairness of the 1 -item general health questionnaire among diverse adolescents”, *Psychological Assessment*, Vol. 29/1, pp. 87-97, <https://doi.org/10.1037/PAS0000323>. [83]
- Brown, S. et al. (2018), “Mental health and reporting bias: Analysis of the GHQ-1 ”, *IZA Discussion Paper* no. 11771, <https://www.iza.org/publications/dp/11771/mental-health-and-reporting-bias-analysis-of-the-ghq-12>. [63]

- Cheng, J. et al. (2016), “Impact of a mass media mental health campaign on psychiatric emergency department visits”, *Canadian Journal of Public Health*, Vol. 107/3, [95]
<https://www.jstor.org/stable/90006480>.
- Coles, M. and S. Coleman (2010), “Barriers to treatment seeking for anxiety disorders: Initial data on the role of mental health literacy”, *Depression and Anxiety*, Vol. 27/1, pp. 63-71, [56]
<https://doi.org/10.1002/DA.20620>.
- Cornelius, B. et al. (2013), “The performance of the K10, K and GHQ-12 to screen for present state DSM-IV disorders among disability claimants”, *BMC Public Health*, Vol. 13/128, [37]
<http://www.biomedcentral.com/1471-2458/13/128>.
- Crawford, M. et al. (2011), “Selecting outcome measures in mental health: The views of service users”, [111]
Journal of Mental Health, Vol. 20/4, pp. 336-346, <https://doi.org/10.3109/09638237.2011.577114>.
- Crockett, L. et al. (2005), “Measurement equivalence of the center for epidemiological studies depression scale for Latino and Anglo adolescents: A national study”, [79]
Journal of Consulting and Clinical Psychology, Vol. 73/1, pp. 47-58, <https://doi.org/10.1037/0022-006X.73.1.47>.
- Dadfar, M. et al. (2018), “Reliability, validity, and factorial structure of the World Health Organization-5 Well-Being Index (WHO-5) in Iranian psychiatric outpatients”, [18]
Trends in Psychiatry and Psychotherapy, Vol. 40/2, pp. 79-84, <https://doi.org/10.1590/2237-6089-2017-0044>.
- de Bienassis, K. et al. (2021), “Patient-reported indicators in mental health care: Towards international standards among members of the OECD”, [139]
International Journal for Quality in Health Care, Vol. 6/34(Suppl 1), <https://doi.org/10.1093/INTQHC/MZAB020>.
- de Graaf, R. et al. (2000), “Psychiatric and sociodemographic predictors of attrition in a longitudinal study the Netherlands Mental Health Survey and Incidence Study (NEMESIS)”, [60]
American Journal of Epidemiology, Vol. 152/11, pp. 1039-1047, <https://doi.org/10.1093/AJE/152.11.1039>.
- de Graaf, R., M. Have and S. Van Dorsselaer (2010), “The Netherlands Mental Health Survey and Incidence Study-2 (NEMESIS-2): Design and methods”, [69]
International Journal of Methods in Psychiatric Research, Vol. 19/3, pp. 125-141, <https://doi.org/10.1002/MPR.317>.
- Deady, M. et al. (2020), “Suicide awareness campaigns: Are they a valid prevention strategy?”, in [93]
What Can Be Done to Decrease suicidal Behaviour in Australia? A Call to Action,
<https://www.blackdoginstitute.org.au/wp-content/uploads/2020/09/What-Can-Be-Done-To-Decrease-Suicide-Chapter-3-Awareness-Campaigns.pdf>.
- Demirchyan, A., V. Petrosyan and M. Thompson (2011), “Psychometric value of the Center for [82]
Epidemiologic Studies Depression (CES-D) scale for screening of depressive symptoms in Armenian population”, *Journal of Affective Disorders*, Vol. 133/3, pp. 489-498,
<https://doi.org/10.1016/j.jad.2011.04.042>.
- Dere, J. et al. (2015), “Cross-cultural examination of measurement invariance of the Beck Depression [149]
Inventory-II”, *Psychological Assessment*, Vol. 27/1, pp. 68-81, <https://doi.org/10.1037/pas0000026>.

- Dewa, L. et al. (2021), “Reflections, impact and recommendations of a co-produced qualitative study with young people who have experience of mental health difficulties”, *Health Expectations*, Vol. 24/S1, pp. 134-146, <https://doi.org/10.1111/HEX.13088>. [125]
- Dhingra, S. et al. (2011), “PHQ-8 days: A measurement option for DSM-5 Major Depressive Disorder (MDD) severity”, *Population Health Metrics*, Vol. 9/11, <https://doi.org/10.1186/1478-7954-9-11>. [42]
- Dowling, N. et al. (2016), “Measurement and control of bias in patient reported outcomes using multidimensional item response theory”, *BMC Medical Research Methodology*, Vol. 16/63, <https://doi.org/10.1186/s12874-016-0161-z>. [150]
- Drapeau, A., R. Boyer and F. Diallo (2011), “Discrepancies between survey and administrative data on the use of mental health services in the general population: Findings from a study conducted in Québec”, *BMC Public Health*, Vol. 11/837, pp. 1-10, <https://doi.org/10.1186/1471-2458-11-837>. [72]
- Dunn, K. et al. (2009), “Quantification and examination of depression-related mental health literacy”, *Journal of Evaluation in Clinical Practice*, Vol. 15/4, pp. 650-653, <https://doi.org/10.1111/J.1365-2753.2008.01067.X>. [55]
- Easton, S. et al. (2017), “The Kessler psychological distress scale: Translation and validation of an Arabic version”, *Health and Quality of Life Outcomes*, Vol. 15/1, <https://doi.org/10.1186/S12955-017-0783-9>. [9]
- Eaton, W. (2004), “Center for Epidemiologic Studies Depression scale: Review and revision (CESD and CESD-R)”, in M. E. Maruish (ed.), *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment: Instruments for Adults*, Lawrence Erlbaum Associates Publishers, <https://psycnet.apa.org/record/2004-14941-011>. [132]
- Eaton, W. et al. (1992), “Psychopathology and attrition in the epidemiologic catchment area surveys”, *American Journal of Epidemiology*, Vol. 135/9, pp. 1051-1059, <https://doi.org/10.1093/OXFORDJOURNALS.AJE.A116399>. [61]
- El-Den, S. et al. (2018), “The psychometric properties of depression screening tools in primary healthcare settings: A systematic review”, *Journal of Affective Disorders*, Vol. 225/1, pp. 503-522, <https://doi.org/10.1016/j.jad.2017.08.060>. [8]
- Elovainio, M. et al. (2020), “General Health Questionnaire (GHQ-12), Beck Depression Inventory (BDI-6), and Mental Health Index (MHI-5): Psychometric and predictive properties in a Finnish population-based sample”, *Psychiatry Research*, Vol. 289, p. 112973, <https://doi.org/10.1016/j.psychres.2020.112973>. [6]
- Epstein, J., P. Barker and L. Kroutil (2001), “Mode effects in self-reported mental health data”, *Public Opinion Quarterly*, Vol. 65/4, pp. 529-549, <https://doi.org/10.1086/323577>. [116]
- Eurofound (2021), *Living, Working and COVID-19: Mental health and trust decline across EU as pandemic enters another year*, Publications Office of the European Union, <https://www.eurofound.europa.eu/topic/covid-19>. [102]
- Eurofound (n.d.), *European Quality of Life Surveys (EQLS) (database)*, <https://www.eurofound.europa.eu/surveys/european-quality-of-life-surveys> (accessed on 10 June 2022). [107]

- Eurostat (n.d.), *European health interview survey (EHIS)*, [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:European_health_interview_survey_\(EHIS\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:European_health_interview_survey_(EHIS)) (accessed on 16 June 2021). [66]
- Eurostat (n.d.), *European Health Interview Survey (EHIS): Reference Metadata in Euro SDMX Metadata Structure (ESMS)*, https://ec.europa.eu/eurostat/cache/metadata/en/hlth_det_esms.htm (accessed on 25 March 2022). [74]
- Eurostat (n.d.), *European Union Statistics on Income and Living Conditions (EU-SILC) (database)*, <https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions> (accessed on 10 June 2022). [65]
- Exton, C., C. Smith and D. Vandendriessche (2015), “Comparing happiness across the world: Does culture matter?”, *OECD Statistics Working Papers*, No. 2015/4, OECD Publishing, Paris, <https://doi.org/10.1787/5jrqpzd9bs2-en>. [92]
- Folkhälsomyndigheten (2022), *Synen på psykisk ohälsa och suicid: En befolkningsundersökning om kunskaper och attityder [The view of mental illness and suicide: A population survey on knowledge and attitudes]*, <https://www.folkhalsomyndigheten.se/publikationer-och-material/publikationsarkiv/s/synen-pa-psykisk-ohalsa-och-suicid-/?pub=105538#105580>. [67]
- Foulkes, L. and J. Andrews (2023), “Are mental health awareness efforts contributing to the rise in reported mental health problems? A call to test the prevalence inflation hypothesis”, *New Ideas in Psychology*, Vol. 69, p. 101010, <https://doi.org/10.1016/J.NEWIDEAPSYCH.2023.101010>. [97]
- Fried, E. (2017), “The 5 symptoms of major depression: Lack of content overlap among seven common depression scales”, *Journal of Affective Disorders*, Vol. 208, pp. 191-197, <https://doi.org/10.1016/j.jad.2016.10.019>. [146]
- Furukawa, T. and D. Goldberg (1999), “Cultural invariance of likelihood ratios for the General Health Questionnaire”, *Lancet*, Vol. 353/9152, pp. 561-562, [https://doi.org/10.1016/S0140-6736\(98\)05470-1](https://doi.org/10.1016/S0140-6736(98)05470-1). [145]
- Furukawa, T. et al. (2001), “Stratum-specific likelihood ratios of two versions of the General Health Questionnaire”, *Psychological Medicine*, Vol. 31/3, pp. 519-529, <https://doi.org/10.1017/s0033291701003713>. [144]
- Furukawa, T. et al. (2003), “The performance of the K and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being”, *Psychological Medicine*, Vol. 33/2, pp. 357-362, <https://doi.org/10.1017/S0033291702006700>. [36]
- Gallup (n.d.), *Gallup World Poll (database)*, <https://www.gallup.com/analytics/318875/global-research.aspx> (accessed on 18 June 2021). [120]
- García, M. et al. (2005), “Comparison between telephone and self-administration of Short Form Health Survey Questionnaire (SF-36).”, *Gaceta sanitaria / S.E.S.P.A.S.*, Vol. 19/6, pp. 433-439, [https://doi.org/10.1016/S0213-9111\(05\)71393-5](https://doi.org/10.1016/S0213-9111(05)71393-5). [143]
- Garland, A. et al. (2018), “Use of the WHO’s perceived Well-Being Index (WHO-5) as an efficient and potentially valid screen for depression in a low income country”, *Families, Systems and Health*, Vol. 36/2, pp. 148-158, <https://doi.org/10.1037/FSH0000344>. [19]

- Gibbons, C. and S. Skevington (2018), “Adjusting for cross-cultural differences in computer-adaptive tests of quality of life”, *Quality of Life Research*, Vol. 27/4, pp. 1027-1039, <https://doi.org/10.1007/s11136-017-1738-7>. [141]
- Gill, S. et al. (2007), “Validity of the mental health component scale of the 12-item Short-Form Health Survey (MCS-12) as measure of common mental disorders in the general population”, *Psychiatry Research*, Vol. 152/1, pp. 63-71, <https://doi.org/10.1016/j.psychres.2006.11.005>. [26]
- Gnambs, T. and K. Kaspar (2017), “Socially desirable responding in web-based questionnaires: A meta-analytic review of the candor hypothesis”, *Assessment*, Vol. 24/6, pp. 746-762, <https://doi.org/10.1177/1073191115624547>. [119]
- Goldberg, D., T. Oldehinkel and J. Ormel (1998), “Why GHQ threshold varies from one place to another”, *Psychol Med*, Vol. 28/4, pp. 915-921, <https://doi.org/10.1017/s0033291798006874>. [151]
- Guo, C. et al. (2015), “Psychometric evaluation of the Mental Health Continuum-Short Form (MHC-SF) in Chinese adolescents: A methodological study”, *Health and Quality of Life Outcomes*, Vol. 13/198, <https://doi.org/10.1186/S12955-015-0394-2>. [46]
- Hancock, N. et al. (2012), “Participation of mental health consumers in research: Training addressed and reliability assessed”, *Australian Occupational Therapy Journal*, Vol. 59/3, pp. 218-224, <https://doi.org/10.1111/J.1440-1630.2012.01011.X>. [124]
- Hanmer, J., R. Hays and D. Fryback (2007), “Mode of administration is important in US national estimates of health-related quality of life”, *Medical Care*, Vol. 45/12, pp. 1171-1179, <https://doi.org/10.1097/MLR.0b013e3181354828>. [142]
- Hapke, U. et al. (2022), “Depressive symptoms in the general population before and in the first year of the COVID-19 pandemic: Results of the GEDA 019/ 0 0 study”, *Journal of Health Monitoring*, Vol. 7/4, <https://doi.org/10.25646/10664>. [158]
- Haroz, E. et al. (2017), “How is depression experienced around the world? A systematic review of qualitative literature”, *Social Science & Medicine*, Vol. 183, pp. 151-162, <https://doi.org/10.1016/j.socscimed.2016.12.030>. [31]
- Harvey, S. (n.d.), *Do mental health awareness campaigns work? Let's look at the evidence*, Black Dog Institute, <https://www.blackdoginstitute.org.au/news/do-mental-health-awareness-campaigns-work-lets-look-at-the-evidence/> (accessed on 25 March 2022). [96]
- Hinshaw, S. and A. Stier (2008), “Stigma as related to mental disorders”, *Annual Review of Clinical Psychology*, Vol. 4, pp. 367-393, <https://doi.org/10.1146/annurev.clinpsy.4.022007.141245>. [53]
- Hoeymans, N. et al. (2004), “Measuring mental health of the Dutch population: A comparison of the GHQ-12 and the MHI-5”, *Health and Quality of Life Outcomes*, Vol. 2/23, <https://doi.org/10.1186/1477-7525-2-2>. [154]
- Holbrook, A., M. Green and J. Krosnick (2003), “Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias”, *Public Opinion Quarterly*, Vol. 67/1, pp. 79-125, <https://doi.org/10.1086/346010>. [118]
- Hollifield, M. et al. (2002), “Measuring trauma and health status in refugees: A critical review”, *JAMA*, Vol. 288/5, pp. 611-621, <https://doi.org/10.1001/jama.288.5.611>. [88]

- Huang, F. et al. (2006), “Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients”, *Journal of General Internal Medicine*, Vol. 21/6, pp. 547-552, <https://doi.org/10.1111/j.1525-1497.2006.00409.x>. [28]
- Huang, W. and S. Wong (2014), “Cross-Cultural Validation”, in *Encyclopedia of Quality of Life and Well-Being Research*, Springer Netherlands, https://doi.org/10.1007/978-94-007-0753-5_630. [87]
- Iwata, N. and S. Buka (2002), “Race/ethnicity and depressive symptoms: A cross-cultural/ethnic comparison among university students in East Asia, North and South America”, *Social Science and Medicine*, Vol. 55/12, pp. 2243-2252, [https://doi.org/10.1016/S0277-9536\(02\)00003-5](https://doi.org/10.1016/S0277-9536(02)00003-5). [81]
- Jacomb, P. et al. (2002), “Predictors of refusal to participate: A longitudinal health survey of the elderly in Australia”, *BMC Public Health*, Vol. 2/4, <https://doi.org/10.1186/1471-2458-2-4>. [68]
- Jacomb, P. et al. (1999), “Emotional response of participants to a mental health survey”, *Social Psychiatry and Psychiatric Epidemiology*, Vol. 34/2, pp. 80-84, <https://doi.org/10.1007/S001270050115>. [110]
- Jin, K. and W. Wang (2014), “Generalized IRT models for extreme response style”, *Educational and Psychological Measurement*, Vol. 74/1, pp. 116-138, <https://doi.org/10.1177/0013164413498876>. [152]
- Jones, N. et al. (2021), “Lived experience, research leadership, and the transformation of mental health services: Building a researcher pipeline”, *Psychiatric Services*, Vol. 72/5, pp. 591-593, <https://doi.org/10.1176/appi.ps.202000468>. [122]
- Jorm, A. et al. (1994), “Do mental health surveys disturb? Further evidence”, *Psychological Medicine*, Vol. 24/1, pp. 233-237, <https://doi.org/10.1017/S0033291700026994>. [109]
- Joshanloo, M. et al. (2013), “Measurement invariance of the Mental Health Continuum-Short Form (MHC-SF) across three cultural groups”, *Personality and Individual Differences*, Vol. 55/7, pp. 755-759, <https://doi.org/10.1016/J.PAID.2013.06.002>. [86]
- Kessler, R. et al. (2002), “Short screening scales to monitor population prevalences and trends in non-specific psychological distress”, *Psychological Medicine*, Vol. 32/6, pp. 959-976, <https://doi.org/10.1017/S0033291702006074>. [89]
- Kessler, R. et al. (2004), “The US National Comorbidity Survey Replication (NCS-R): Design and field procedures”, *International Journal of methods in Psychiatric Research*, Vol. 13/2, pp. 69-92, <https://doi.org/10.1002/MPR.167>. [64]
- Knudsen, A. et al. (2021), “Prevalence of mental disorders, suicidal ideation and suicides in the general population before and during the COVID-19 pandemic in Norway: A population-based repeated cross-sectional analysis”, *The Lancet Regional Health Europe*, Vol. 4, <https://doi.org/10.1016/J.LANEPE.2021.100071>. [160]
- Kroenke, K., R. Spitzer and J. Williams (2001), “The PHQ-9: Validity of a brief depression severity measure”, *Journal of General Internal Medicine*, Vol. 16/9, <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>. [137]

- Kroenke, K. et al. (2009), “An ultra-brief screening scale for anxiety and depression: The PHQ-4”, [12]
Psychosomatics, Vol. 50/6, pp. 613-621, <https://doi.org/10.1176/appi.psy.50.6.613>.
- Kroenke, K. et al. (2007), “Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and [44]
 detection”, *Annals of Internal Medicine*, Vol. 146/5, pp. 317-325,
<https://doi.org/10.7326/0003-4819-146-5-200703060-00004>.
- Kroenke, K. et al. (2008), “The PHQ-8 as a measure of current depression in the general population”, [136]
Journal of Affective Disorders, Vol. 114/1-3, pp. 163-173, <https://doi.org/10.1016/j.jad.2008.06.026>.
- Krumpal, I. (2013), “Determinants of social desirability bias in sensitive surveys: A literature review”, [59]
Quality and Quantity, Vol. 47/4, pp. 2025-2047, <https://doi.org/10.1007/S11135-011-9640-9>.
- Labott, S. et al. (2013), “Emotional risks to respondents in survey research: Some empirical evidence”, [108]
Journal of Empirical Research on Human Research Ethics, Vol. 8/4, p. 53,
<https://doi.org/10.1525/JER.2013.8.4.53>.
- Lakeman, R. and M. FitzGerald (2009), “The ethics of suicide research”, *Crisis*, Vol. 30/1, pp. 13-19, [128]
<https://doi.org/10.1027/0227-5910.30.1.13>.
- Lamers, S. et al. (2011), “Evaluating the psychometric properties of the Mental Health Continuum-Short [20]
 Form (MHC-SF)”, *Journal of Clinical Psychology*, Vol. 67/1, pp. 99-110,
<https://doi.org/10.1002/jclp.20741>.
- Latkin, C. et al. (2017), “The relationship between social desirability bias and self-reports of health, [114]
 substance use, and social network factors among urban substance users in Baltimore, Maryland”,
Addictive Behaviors, Vol. 73, pp. 133-136, <https://doi.org/10.1016/J.ADDBEH.2017.05.005>.
- Lee, J. et al. (2010), “Cross-cultural considerations in administering the Center for Epidemiologic [84]
 Studies Depression Scale”, *Gerontology*, Vol. 57/5, pp. 455-461, <https://doi.org/10.1159/000318030>.
- Leong, F., P. Priscilla Lui and Z. Kalibatseva (2019), “Multicultural Issues in Clinical Psychological [23]
 Assessment”, in Selbom, M. and J. Suhr (eds.), *The Cambridge Handbook of Clinical Assessment
 and Diagnosis*, Cambridge University Press, <https://doi.org/10.1017/9781108235433>.
- Löwe, B. et al. (2010), “A -item measure of depression and anxiety: Validation and standardization of [13]
 the Patient Health Questionnaire-4 (PHQ-4) in the general population”, *Journal of Affective
 Disorders*, Vol. 122/1-2, pp. 86-95, <https://doi.org/10.1016/j.jad.2009.06.019>.
- Lowthian, P. and L. Lloyd (2020), *How face-to-face interviewer attitudes and beliefs moderate the [162]
 effect of monetary incentive on UK Labour Force Survey response rates*, Office for National
 Statistics (ONS),
<https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/methodologies/howfacetofaceinterviewerattitudesandbeliefsmoderatetheeffectofmonetaryincentiveonuklabourforcesurveyresponserates>.

- Manea, L., S. Gilbody and D. McMillan (2015), “A diagnostic meta-analysis of the Patient Health Questionnaire-9 (PHQ-9) algorithm scoring method as a screen for depression”, *General Hospital Psychiatry*, Vol. 37/1, pp. 67-75, <https://doi.org/10.1016/j.genhosppsych.2014.09.009>. [131]
- Mauz, E. et al. (2022), “Time trends of mental health indicators in Germany’s adult population before and during the COVID-19 pandemic”, *medRxiv*, <https://doi.org/10.1101/2022.10.09.22280826>. [159]
- Medina-Mora, M. et al. (2008), “The Mexican National Comorbidity Survey (M-NCS): Overview and results”, in Kessler, R. and T. Üstün (eds.), *The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders*, Cambridge University Press, <https://digitallibrary.un.org/record/700329?ln=en>. [153]
- Moriarty, D., M. Zack and R. Kobau (2003), “The Centers for Disease Control and Prevention’s healthy days measures: Population tracking of perceived physical and mental health over time”, *Health and Quality of Life Outcomes*, Vol. 1/37, <https://doi.org/10.1186/1477-7525-1-37>. [21]
- Mostafa, T. et al. (2021), “Missing at random assumption made more plausible: Evidence from the 195 British birth cohort”, *Journal of Clinical Epidemiology*, Vol. 136, pp. 44-54, <https://doi.org/10.1016/J.JCLINEPI.2021.02.019>. [62]
- Moum, T. (1998), “Mode of administration and interviewer effects in self-reported symptoms of anxiety and depression”, *Social Indicators Research*, Vol. 45/1-3, pp. 279-318, <https://doi.org/10.1023/a:1006958100504>. [115]
- Murphy, R. and B. Hallahan (2016), “Differences between DSM-IV and DSM-5 as applied to general adult psychiatry”, *Irish Journal of Psychological Medicine*, Vol. 33/3, pp. 135-141, <https://doi.org/10.1017/IPM.2015.54>. [30]
- New Zealand Government (2018), *He Ara Oranga : Report of the Government Inquiry into Mental Health and Addiction*, Mental Health and Addiction Inquiry, <https://mentalhealth.inquiry.govt.nz/inquiry-report/he-ara-oranga/>. [34]
- NHS Health Scotland (2016), *Warwick-Edinburgh Mental Well-being Scale (WEMWBS): User guide-Version 2*, NHS Health Scotland, Edinburgh, <https://s3.amazonaws.com/helpscout.net/docs/assets/5f97128852faff0016af3a34/attachment/s/5fe10a9eb624c71b7985b8f3/WEMWBS-Scale.pdf>. [15]
- NHS Health Scotland (2008), *Review of scales of positive mental health validated for use with adults in the UK: Technical report*, NHS Health Scotland, Edinburgh, <http://www.healthscotland.scot/media/2244/review-of-scales-of-positive-mental-health-validated-for-use-with-adults-in-the-uk.pdf>. [3]
- O’Connor, D. and R. Parslow (2010), “Mental health scales and psychiatric diagnoses: Responses to GHQ-12, K-10 and CIDI across the lifespan”, *Journal of Affective Disorders*, Vol. 121/3, pp. 263-267, <https://doi.org/10.1016/j.jad.2009.06.038>. [27]
- OECD (2021), *A New Benchmark for Mental Health Systems: Tackling the Social and Economic Costs of Mental Ill-Health*, OECD Health Policy Studies, OECD Publishing, Paris, <https://doi.org/10.1787/4ed890f6-en>. [71]

- OECD (2021), *COVID-19 and Well-being: Life in the Pandemic*, OECD Publishing, Paris, [99]
<https://doi.org/10.1787/1e1ecb53-en>.
- OECD (2021), *Fitter Minds, Fitter Jobs: From Awareness to Change in Integrated Mental Health, Skills and Work Policies*, Mental Health and Work, OECD Publishing, [57]
<https://doi.org/10.1787/a0815d0f-en>.
- OECD (2021), “Tackling the mental health impact of the COVID-19 crisis: An integrated, whole-of-society response”, *OECD Policy Responses to Coronavirus (COVID-19)*, [98]
<https://www.oecd.org/coronavirus/policy-responses/tackling-the-mental-health-impact-of-the-covid-19-crisis-an-integrated-whole-of-society-response-0ccaafa0b/>.
- OECD (2020), *How's Life? 2020: Measuring Well-being*, OECD Publishing, Paris, [140]
<https://doi.org/10.1787/9870c393-en>.
- OECD (2017), *OECD Guidelines on Measuring Trust*, OECD Publishing, Paris, [2]
<https://doi.org/10.1787/9789264278219-en>.
- OECD (2013), *OECD Guidelines on Measuring Subjective Well-being*, OECD Publishing, Paris, [1]
<https://doi.org/10.1787/9789264191655-en>.
- Ohno, S. et al. (2017), “Smallest detectable change and test-retest reliability of a self-reported outcome measure: Results of the Center for Epidemiologic Studies Depression Scale, General Self-Efficacy Scale, and 12-item General Health Questionnaire”, *Journal of Evaluation in Clinical Practice*, Vol. 23/6, pp. 1348-1354, <https://doi.org/10.1111/JEP.12795>. [5]
- Parkerson, H. et al. (2015), “Cultural-based biases of the GAD-7”, *Journal of Anxiety Disorders*, Vol. 31, pp. 38-42, <https://doi.org/10.1016/j.janxdis.2015.01.005>. [85]
- Paulhus, D. (1984), “Two-component models of socially desirable responding”, *Journal of Personality and Social Psychology*, Vol. 46/3, pp. 598-609, [112]
<https://psycnet.apa.org/doi/10.1037/0022-3514.46.3.598>.
- Pedrelli, P. et al. (2013), “Reliability and validity of the Symptoms of Depression Questionnaire (SDQ)”, *CNS Spectrums*, Vol. 19/6, pp. 535-546, <https://doi.org/10.1017/S1092852914000406>. [147]
- Petrillo, G. et al. (2015), “The Mental Health Continuum Short Form (MHCSF) as a measure of well-being in the Italian context”, *Social Indicators Research*, Vol. 121, pp. 291-312, [47]
<https://doi.org/10.1007/s11205-014-0629-3>.
- Pettersson, A. et al. (2015), “Which instruments to support diagnosis of depression have sufficient accuracy? A systematic review”, *Nordic Journal of Psychiatry*, Vol. 69/7, pp. 497-508, [41]
<https://doi.org/10.3109/08039488.2015.1008568>.
- Pintea, S. and R. Moldovan (2009), “The Receiver-Operating Characteristic (ROC) analysis: Fundamentals and applications in clinical psychology”, *Journal of Cognitive and Behavioral Psychotherapies*, Vol. 9/1, pp. 49-66, <https://psycnet.apa.org/record/2009-04815-004> (accessed on 0). [24]

- Ploubidis, G., E. McElroy and H. Moreira (2019), “A longitudinal examination of the measurement equivalence of mental health assessments in two British birth cohorts”, *Longitudinal and Life Course Studies*, Vol. 10/4, pp. 471-489, <https://doi.org/10.1332/175795919X15683588979486>. [105]
- Posner, S. et al. (2001), “Factor variability of the Center for Epidemiological Studies Depression Scale (CES-D) among urban Latinos”, *Ethnicity and Health*, Vol. 6/2, pp. 137-144, <https://doi.org/10.1080/13557850120068469>. [80]
- Presser, S. and L. Stinson (1998), “Data collection mode and social desirability bias in self-reported religious attendance”, *American Sociological Review*, Vol. 63/1, pp. 137-145, <https://doi.org/10.2307/2657486>. [113]
- Public Health Agency of Sweden (2022), *Ny nationell strategi för psykisk hälsa och suicidprevention*, <https://www.folkhalsomyndigheten.se/livsvillkor-levnadsvanor/psykisk-halsa-och-suicidprevention/psykisk-halsa/nationell-strategi/>. [35]
- Richardson, L. et al. (2010), “Evaluation of the Patient Health Questionnaire-9 item for detecting major depression among adolescents”, *Pediatrics*, Vol. 126/6, pp. 1117-1123, <https://doi.org/10.1542/peds.2010-0852>. [77]
- Rivera-Riquelme, M., J. Piqueras and P. Cuijpers (2019), “The revised Mental Health Inventory-5 (MHI-5) as an ultra-brief screening measure of bidimensional mental health in children and adolescents”, *Psychiatry Research*, Vol. 274, pp. 247-253, <https://doi.org/10.1016/J.PSYCHRES.2019.02.045>. [40]
- Rose, M. and J. Devine (2014), “Assessment of patient-reported symptoms of anxiety”, *Dialogues in Clinical Neuroscience*, Vol. 16/2, pp. 197-211, <https://doi.org/10.31887/DCNS.2014.16.2/mrose>. [43]
- Rumpf, H. et al. (2001), “Screening for mental health: Validity of the MHI-5 using DSM-IV Axis I psychiatric disorders as gold standard”, *Psychiatry Research*, Vol. 105/3, pp. 243-253, [https://doi.org/10.1016/S0165-1781\(01\)00329-8](https://doi.org/10.1016/S0165-1781(01)00329-8). [135]
- Santini, Z. et al. (2022), “Higher levels of mental wellbeing predict lower risk of common mental disorders in the Danish general population”, *Mental Health & Prevention*, Vol. 26, p. 200233, <https://doi.org/10.1016/j.mhp.2022.200233>. [148]
- Santini, Z. et al. (2020), “Measuring positive mental health and flourishing in Denmark: Validation of the mental health continuum-short form (MHC-SF) and cross-cultural comparison across three countries”, *Health and Quality of Life Outcomes*, Vol. 18/297, <https://doi.org/10.1186/s12955-020-01546-2>. [48]
- Santomauro, D. et al. (2021), “Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic”, *The Lancet*, Vol. 398/10312, pp. 1700-1712, [https://doi.org/10.1016/S0140-6736\(21\)02143-7](https://doi.org/10.1016/S0140-6736(21)02143-7). [161]
- Santor, D. and J. Coyne (1997), “Shortening the CES-D to improve its ability to detect cases of depression”, *Psychological Assessment*, Vol. 9/3, pp. 233-243, <https://doi.org/10.1037/1040-3590.9.3.233>. [133]

- Schmitz, N., J. Kruse and W. Tress (2001), “Improving screening for mental disorders in the primary care setting by combining the GHQ-12 and SCL-90-R subscales”, *Comprehensive Psychiatry*, Vol. 42/2, pp. 166-173, <https://doi.org/10.1053/COMP.2001.19751>. [4]
- Scholz, B. et al. (2021), “‘People just need to try it to be converted!’: A picture of consumer mental health research in Australia and New Zealand”, *Issues in Mental Health Nursing*, Vol. 42/3, pp. 249-255, <https://doi.org/10.1080/01612840.2020.1795763>. [127]
- Shah, N. et al. (2021), “Short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS): Performance in a clinical sample in relation to PHQ-9 and GAD-7”, *Health and Quality of Life Outcomes*, Vol. 19/260, <https://doi.org/10.1186/s12955-021-01882-x>. [17]
- Shah, N. et al. (2018), “Responsiveness of the Short Warwick Edinburgh Mental Well-Being Scale (SWEMWBS): Evaluation a clinical sample”, *Health and Quality of Life Outcomes*, Vol. 16/1, <https://doi.org/10.1186/s12955-018-1060-2>. [104]
- Siegerink, V., M. Shinwell and Ž. Žarnić (2022), “Measuring the non-financial performance of firms through the lens of the OECD Well-being Framework: A common measurement framework for ‘Scope 1’ Social performance”, *OECD Papers on Well-being and Inequalities*, No. 3, OECD Publishing, Paris, https://www.oecd-ilibrary.org/social-issues-migration-health/measuring-the-non-financial-performance-of-firms-through-the-lens-of-the-oecd-well-being-framework_28850c7f-en (accessed on 18 May 2022). [138]
- Sigmon, S. et al. (2005), “Gender differences in self-reports of depression: The response bias hypothesis revisited”, *Sex Roles*, Vol. 53/5-6, pp. 401-411, <https://doi.org/10.1007/s11199-005-6762-3>. [75]
- Singer, E., H. Hippler and N. Schwarz (1992), “Confidentiality assurances in surveys: Reassurance or threat?”, *International Journal of Public Opinion Research*, Vol. 4/3, pp. 256-268, <https://doi.org/10.1093/IJPOR/4.3.256>. [156]
- Singer, E., D. Von Thurn and E. Miller (1995), “Confidentiality assurances and response: A quantitative review of the experimental literature”, *Public Opinion Quarterly*, Vol. 59/1, pp. 66-77, <https://doi.org/10.1086/269458>. [58]
- Smith, D. and B. Oreskes (2019), “Homeless population’s mental illness, substance abuse under-reported”, *Los Angeles Times*, <https://www.latimes.com/california/story/2019-10-07/homeless-population-mental-illness-disability>. [157]
- Spitzer, R. et al. (2006), “A brief measure for assessing generalized anxiety disorder: The GAD-7”, *Archives of Internal Medicine*, Vol. 166/10, pp. 1092-1097, <https://doi.org/10.1001/ARCHINTE.166.10.1092>. [11]
- Statistics Canada (2021), *Measuring Population Mental Health in Canada*, <https://www.slideshare.net/StatsCommunications/oecd-wellbeing-and-mental-health-conference-jennifer-ali-statcan>. [29]
- Stewart-Brown, S. (2021), *15 years on: Insights and reflections on the Warwick-Edinburgh Mental Wellbeing Scales (WEMWBS)*, What Works Wellbeing, <https://whatworkswellbeing.org/resources/insights-and-reflections-on-the-warwick-edinburgh-mental-wellbeing-scales-wemwbs/>. [16]

- Stewart-Brown, S. et al. (2009), “Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a Rasch analysis using data from the Scottish Health Education Population Survey”, *Health and Quality of Life Outcomes*, Vol. 7/1, p. 15, <https://doi.org/10.1186/1477-7525-7-15>. [45]
- St-Pierre, M. and Y. Béland (2004), “Mode effects in the Canadian Community Health Survey: A comparison of CAPI and CATI”, in *2004 Proceedings of the American Statistical Association Meeting, Survey Research Methods*, American Statistical Association, Toronto, Canada, https://www.statcan.gc.ca/en/statistical-programs/document/3226_D16_T9_V5. [117]
- Strand, B. et al. (2003), “Measuring the mental health status of the Norwegian population: A comparison of the instruments SCL-25, SCL-10, SCL-5 and MHI-5 (SF-36)”, *Nordic Journal of Psychiatry*, Vol. 57/2, pp. 113-118, <https://doi.org/10.1080/08039480310000932>. [7]
- Streiner, D. and J. Cairney (2007), “What’s under the ROC? An introduction to Receiver Operating Characteristics curves”, *The Canadian Journal of Psychiatry*, Vol. 52/2, pp. 121-128, <https://doi.org/10.1177/070674370705200210> (accessed on 18 May 2022). [25]
- Suhr, D. (2006), *Exploratory or Confirmatory Factor Analysis?*, SAS Users Group International (SUGI), San Francisco, CA, <https://support.sas.com/resources/papers/proceedings/proceedings/sugi31/200-31.pdf>. [22]
- Sunderland, M. et al. (2019), “Self-Report Scales for Common Mental Disorders”, in *The Cambridge Handbook of Clinical Assessment and Diagnosis*, Cambridge University Press, <https://doi.org/10.1017/9781108235433.019>. [33]
- Tambling, R., C. D’Aniello and B. Russell (2021), “Mental health literacy: A critical target for narrowing racial disparities in behavioral health”, *International Journal of Mental Health and Addiction*, <https://doi.org/10.1007/s11469-021-00694-w>. [54]
- Taylor, A. et al. (2011), “Methodological issues associated with collecting sensitive information over the telephone: Experience from an Australian non-suicidal self-injury (NSSI) prevalence study”, *BMC Medical Research Methodology*, Vol. 11, <https://doi.org/10.1186/1471-2288-11-20>. [129]
- Tennant, R. et al. (2007), “The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation”, *Health and Quality of Life Outcomes*, Vol. 5/63, <https://doi.org/10.1186/1477-7525-5-63>. [14]
- Thorsen, S. et al. (2013), “The predictive value of mental health for long-term sickness absence: the Major Depression Inventory (MDI) and the Mental Health Inventory (MHI-5) compared”, *BMC Medical Research Methodology*, Vol. 13/115, <https://doi.org/10.1186/1471-2288-13-115>. [130]
- Thygesen, L. et al. (2021), “Decreasing mental well-being during the COVID-19 pandemic: A longitudinal study among Danes before and during the pandemic”, *Journal of Psychiatric Research*, Vol. 144, pp. 151-157, <https://doi.org/10.1016/J.JPSYCHIRES.2021.09.035>. [100]
- Topp, C. et al. (2015), “The WHO-5 well-being index: A systematic review of the literature”, *Psychotherapy and Psychosomatics*, Vol. 84/3, pp. 167-176, <https://doi.org/10.1159/000376585>. [49]

- University of Essex, I. (2022), *Understanding Society: Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009 (database)*, 5th Edition. UK Data Service., [70]
<https://doi.org/10.5255/UKDA-SN-6614-16> (accessed on 10 June 2022).
- University of Michigan (2021), *Panel Study of Income Dynamics (database)*, Survey Research Center, [76]
 Institute for Social Research, University of Michigan, Ann Arbor, MI,
<https://psidonline.isr.umich.edu/default.aspx> (accessed on 10 June 2022).
- Vaughn-Coaxum, R., P. Mair and J. Weisz (2016), “Racial/ethnic differences in youth depression [155]
 indicators”, *Clinical Psychological Science*, Vol. 4/2, pp. 239-253,
<https://doi.org/10.1177/2167702615591768>.
- Vilagut, G. et al. (2013), “The Mental Component of the Short-Form 12 Health Survey (SF-12) as a [91]
 measure of depressive disorders in the general population: Results with three alternative scoring
 methods”, *Value in Health*, Vol. 16/4, pp. 564-573, <https://doi.org/10.1016/j.jval.2013.01.006>.
- Vinko, M. et al. (2022), “Positive mental health in Slovenia before and during the COVID-19 [103]
 pandemic”, *Frontiers in Public Health*, Vol. 10, p. 3719, <https://doi.org/10.3389/fpubh.2022.963545>.
- Vistisen, H. et al. (2022), “The less depressive state of Denmark following the second wave of the [101]
 COVID-19 pandemic”, *Acta Neuropsychiatrica*, Vol. 34/3, pp. 163-166,
<https://doi.org/10.1017/NEU.2022.1>.
- Walsh, D. and J. Foster (2021), “A call to action: A critical review of mental health related anti- [94]
 stigma campaigns”, *Frontiers in Public Health*, Vol. 8, p. 990,
<https://doi.org/10.3389/FPUBH.2020.569539/BIBTEX>.
- Ware, J. et al. (2002), *How to score SF-12 items: How to score version 2 of the SF-12 Health Survey*, [90]
<https://www.researchgate.net/publication/291994160>.
- Warr, D., R. Mann and T. Tacticos (2011), “Using peer-interviewing methods to explore place-based [126]
 disadvantage: Dissolving the distance between suits and civilians”, *International Journal of Social
 Research Methodology*, Vol. 14/5, pp. 337-352, <https://doi.org/10.1080/13645579.2010.537527>.
- Warwick Medical School (2021), *The Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS)*, [78]
<https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/>.
- Woodhead, C. et al. (2012), “Impact of exposure to combat during deployment to Iraq and Afghanistan [106]
 on mental health by gender”, *Psychological Medicine*, Vol. 42/9, pp. 1985-1996,
<https://doi.org/10.1017/S003329171100290X>.
- Yamazaki, S., S. Fukuhara and J. Green (2005), “Usefulness of five-item and three-item Mental Health [39]
 Inventories to screen for depressive symptoms in the general population of Japan”, *Health and
 Quality of Life Outcomes*, Vol. 3/1, pp. 1-7, <https://doi.org/10.1186/1477-7525-3-48>.

주

- ¹ 일부 선별 도구에는 경험한 증상과 관련된 문항이 포함되어 있다. ‘선별 도구’는 정신건강 상태의 임상적 진단에 대해 검증된 도구인 반면, ‘경험한 증상’은 응답자에게 ‘현재 PTSD 증상을 경험하고 있는지’ 또는 ‘만성 불안을 겪고 있는지’에 대한 일반적인 문항으로 구성된 비표준화된 형식이라는 점에서 ‘선별 도구’와 ‘경험한 증상’의 범주를 구분한다(표 2.7 참조). 다양한 도구 유형에 대한 자세한 논의는 본 장을 참조한다.
- ² 본 장에서는 SF-12, WHO-5, MHC-SF 및 (S)WEMWBS 등 네 가지 긍정적 정신건강 복합 척도만을 다룬다. 긍정적 정신건강과 주관적 웰빙의 삶의 평가, 정서 및 자아실현적 측면에 대한 측정 지침은 OECD(2013_[11])에서 자세히 다룬다.
- ³ 구성상 항목 수가 적은 선별 도구는 크론바흐 알파 값이 낮다 (특히 크론바흐 알파는 척도에 포함된 총 항목 수에 따른 함수라는 점을 유의해야 함). 이는 단일 검증에 높은 비중을 두기보다는 통계적 품질의 모든 측면에 가중치를 부여하는 것이 중요하다는 점이 재차 강조된다.
- ⁴ GAD-7과 GAD-2의 수렴 타당도를 검증한 다른 두 가지 불안 척도는 Beck 불안 척도(Beck Anxiety Inventory, BAI)와 정신진단검사(Symptom Checklist-90, SCL-90)의 하위 불안 척도이다.
- ⁵ 라쉬 분석은 계량심리 모형을 사용하여 범주형 데이터를 분석하고 잠재된 태도 또는 특성을 식별하고 측정한다.
- ⁶ 정신건강에 대한 낙인과 낮은 수준의 정신건강 이해도가 정신건강 설문조사 항목의 무응답률을 높이는 핵심적인 요인이지만, 낮은 수준의 기관 신뢰, 참여 동기 부재 또는 참여 시간 부족, 언어 장벽, 건강 악화 등의 다른 요인도 낮은 응답률에 영향을 미칠 수 있다(Lowthian 및 Lloyd, 2020_[162]).
- ⁷ 강력한 기밀성을 보장하면 민감한 주제에 대한 무응답률을 낮출 수 있다(Singer, Von Thurn 및 Miller, 1995_[58]). 그러나 응답자가 데이터 기밀성에 대한 상세한 설명을 들은 후 면담에서 위협적이거나 민감한 질문을 예상하고 응답을 미룰 수 있기 때문에 민감하지 않은 주제에 대한 무응답률이 오히려 증가할 수 있다(Singer, Hippler 및 Schwarz, 1992_[156]).
- ⁸ 낙인 방지 지표를 통해 측정한 우울장애 위험과 낙인 간의 상관관계, 즉 정신장애 치료를 받는 것이 바람직하다고 동의하는 사람의 비율과 정신질환은 다른 질병과 동일한 질병이라는 것에 동의하는 사람의 비율은 정신건강에 대한 편견이 적은 국가에서 유병률이 더 낮은 역의 관계를 보여준다. 그러나 이러한 상관관계는 유의미하지 않다.
- ⁹ 국가 간 및 집단 간 비교 가능성은 사소한 측정 문제가 아니며, 과거 일부 OECD 연구에서 인구의 20%가 주어진 12개월 동안 어떤 형태의 정신장애를 경험한다는 역학 조사 결과에 근거하여 인구의 최저 5분위수를 정신적 고통 위험군으로 지정하여 국가 간 비교 문제를 다루었다(OECD, 2021_[57]). 그러나 이러한 접근 방식은 정의상 일정한 유병률을 할당하며, 특히 코로나19 팬데믹의 여파로 OECD 각국 정부가 국민의 정신적 고통, 우울증, 불안, 스트레스 급증에 대처하기 위해 고군분투하고 있는 상황에서 이러한 접근 방식에는 한계가 있다.

¹⁰ WHO-5가 설문조사에 사용된 지리적 범위는 아프리카(알제리, 남아프리카), 아시아(방글라데시, 중국, 인도, 일본, 한국, 스리랑카, 대만, 태국), 유럽(북부, 남부, 동부, 서부 및 중부 유럽), 아메리카(캐나다, 미국, 브라질, 멕시코), 중동(이스라엘, 이란, 레바논) 및 오세아니아(호주, 뉴질랜드)를 포함한다(Topp 외, 2015^[49]).

¹¹ 이러한 새로운 방법은 정신건강 척도의 타당도, 정확성, 비교 가능성 및 효율성을 개선하기 위해 문항 반응이론(IRT)과 같은 현대적인 심리 측정 방법의 적용에 크게 의존하며, 이는 문화 간 차이에 대한 고급 분석에서 상당한 가능성을 보여주었다. 이러한 방법은 IRT 기반 차별기능문항(differential item functioning) 및 문항 고정 또는 동등화의 사용을 통해 유의한 편향을 조정할 수 있다(Dere 외, 2015^[149], Gibbons 및 Skevington, 2018^[141], Vaughn-Coaxum, Mair 및 Weisz, 2016^[155]). 이와 유사하게 극단적인 응답 방식을 기존 방법보다 효과적으로 추정 및 보정하고, 특정 척도에서 극단적인 응답 경향을 정량화할 수 있는 새로운 IRT 모형도 등장했다(Dowling 외, 2016^[150], Jin 및 Wang, 2014^[152]). 이러한 새로운 방법에는 문항 은행(item banking), 적응 검증(adapting testing), 데이터 기반 단축형 척도(data-driven short scales) 및 척도 동등화(scale equating)가 포함된다.

¹² 대부분의 국제 연구에서 이러한 정신질환의 증가 추세가 확인되었지만(OECD, 2021^[98], Santomauro 외, 2021^[161]) 개별 국가에서 정신건강 결과의 궤적이 다소 상이한 것으로 나타났다. 독일의 한 연구에 따르면 팬데믹 발발 첫해에는 우울증 유병률이 감소했지만 2020년 10월에는 증가하기 시작하여 2021년과 2022년에 걸쳐 더욱 증가했다(Hapke 외, 2022^[158], Mauz 외, 2022^[159]). 노르웨이의 한 역학 조사에 따르면 팬데믹 발발 초기(2020년 5월)에 정신장애 유병률이 다소 감소했다가 2020년 9월에는 팬데믹 이전 수준으로 회복되어 정신장애가 비교적 안정적인 수준임을 시사했다(Knudsen 외, 2021^[160]). 이는 2020년 초 65건의 연구를 메타 분석한 결과와 일치하는 것으로, 2020년 3월과 4월에는 정신건강 증상이 평균적으로 소폭 증가했으나 7월에는 감소한 것으로 나타났다. 두 연구 모두 2020년 3분기 초에 종료되었기 때문에 연구를 확장하면 팬데믹 초기에는 정신적 고통에 대한 변화가 거의 없었지만, 2020년 말과 2021년에는 고통이 증가한 독일과 유사한 결과가 도출될 가능성이 열려있다.

¹³ 이러한 패턴은 신체건강 결과에도 존재한다. 미국과 캐나다의 공동 연구에 따르면 자기보고 설문조사 응답자는 면담자가 실시하는 전화 설문조사 응답자보다 건강관련 삶의 질(HRQoL)이 낮다고 보고할 가능성이 더 높았으며(Hanmer, Hays 및 Fryback, 2007^[142]), 스페인의 또 다른 연구에 따르면 면담자가 실시하는 설문조사는 SF-36을 사용하여 측정할 경우 응답자가 신체건강 보다 양호하다고 응답하는 것으로 나타났다(García 외, 2005^[143]).

¹⁴ CCHS 설문조사에는 컴퓨터를 이용한 개별 면접조사(computer-assisted personal interviews, CAPI)와 컴퓨터를 이용한 전화조사(computer-assisted telephone interviews, CATI)가 모두 포함된다. 2001년~2003년에 실시된 설문조사에서는 CATI와 CAPI 면담의 비율을 변경하여 모드 변경이 CCHS 데이터 수집 라운드별 비교 가능성에 미치는 영향에 대해 연구했다. 연구를 통해 모드 변경에 따른 건강 지표 결과의 차이가 발견되었는데, 대면 조사 응답자의 경우 더 높은 비만율을 보고했고 비활동적, 흡연, 의사와의 접촉을 보고할 가능성이 더 높은 것으로 나타났다. 그러나 자기보고 정신건강의 경우 모드 변경에 따른 차이가 나타나지 않았다(St-Pierre 및 Béland, 2004^[117]).

¹⁵ 2020년과 2021년에는 대면 조사를 통해 데이터를 수집하는 많은 국가에서 코로나 19 팬데믹의 영향으로 전화 조사로 전환했다. 이 기간에 정신건강 결과는 글로벌 팬데믹의 영향을 크게 받을 수 있으므로 이러한 모드 전환은 수치에 반영되지 않았다.

¹⁶ 그러나 모든 변경의 원인이 되었던 사회정치적 사건에서 모드 효과를 분리하는 것은 불가능하며, 이는 근본적인 정신건강에 영향을 미쳤을 것으로 예상된다. 본 그림에서 강조된 여러 모드 전환은 심각한 정치적 혼란이나 폭력 사건을 경험한 국가에서 발생했으며, 이는 Gallup이 애초에 데이터 수집 모드를 변경하는 데 영향을 미쳤을 가능성이 높다. 예를 들어 튀르키예의 모드 전환은 2016년 쿠데타 시도가 있었던 시기와 일치하고, 2013년 이라크의 모드 전환은 이 지역에서 ISIS의 활동이 활발해진 시기와 일치한다. 리비아의 모드 전환은 제2차 리비아 내전이 시작된 시기와 일치하는 등 여러 가지 사건이 발생했다. 이러한 사건은 인구의 부정적 정서 균형에 실질적인 영향을 미치며, 모드 효과와는 별개로 그림에 표시된 일부 변화를 유발했을 가능성이 높다.

¹⁷ 예를 들어 멕시코에서는 정신장애가 여전히 수치스러운 질병으로 인식되고 있지만, 멕시코 병존질환조사 (Mexican National Comorbidity Survey) 응답 거부율이 거의 없었고 응답자는 면담자와의 대화 과정에서 처음으로 자신의 정신건강에 대해 거리낌 없이 털어놓는 사례가 자주 발견되었다(Medina-Mora 외, 2008_[153]).

¹⁸ 바닥효과는 척도의 하단에 분포가 몰려 있을 경우 발생하는 반면, 천장효과는 척도의 상단에 분포가 몰려 있을 경우 발생한다. 그림 3.9에서 GHQ-12는 대부분의 응답자가 척도의 하단에 분포되어 있다는 점에서 바닥효과를 보여 준다. 이는 응답자가 정신적 고통에 대한 위험이 심각하지 않음을 나타낸다. 이 척도는 정신적 고통을 경험하는 대상에게 초점을 맞추고 있기 때문에 근본적으로 긍정적 정신건강 수준이 높은 개인을 구분하는 데 있어 민감도가 낮을 수 있다.

¹⁹ 그러나 일부 연구에서는 (S)WEMWBS에 천장효과가 존재함을 시사한다.

²⁰ 엄격한 임상 검증 과정에도 불구하고 임계점수에 대한 비판적인 시각이 여전히 남아 있다. 민감도와 특이도를 최적화하는 준거점수는 인구집단별로 상당히 다를 수 있으며 그 결과 준거점수 사용에 대한 대안 방안이 제시되었다(Goldberg, Oldehinkel 및 Ormel, 1998_[151]). 제안된 여러 방안 중 한 가지는 고정된 임계 값이 아닌 층화우도비(stratum-specific likelihood ratios)를 적용함으로써 보다 세부적인 분류 시스템을 마련 하는 것이다(Furukawa 외, 2001_[144], Furukawa 및 Goldberg, 1999_[145]). 또한 자기보고 증상에 대한 연구에서 새롭게 밝혀진 바에 따르면 주요 우울증 위험을 추정하기 위해 단일 합산 점수와 임상적 준거점수를 사용하면 우울증 연구에 대한 중요한 임상적 결과를 도출하기 어려울 수 있다는 사실을 발견했다(Fried, 2017_[146]). 해당 문제를 극복하고자 일부 연구자들은 강력하고 일반화할 수 있는 결론을 도출하기 위해 여러 우울증 척도를 사용하거나 중요한 DSM 증상에 포함되지 않은 척도(예: 우울증 증상 설문지)를 사용할 것을 권장했다(Pedrelli 외, 2013_[147]). 이러한 점에 주목하는 것은 유용하지만 임상 수준의 의료 전문가가 아닌 거시적 수준에서 인구 정신건강을 측정하는 정부 기관의 경우, 임계점수의 사용이 부적절하거나 충분한 정보를 제공하지 않는다고 판단할 근거가 거의 없다.

²¹ 다른 연구에서는 54점~76점 범위의 준거점수를 사용했다(Thorsen 외, 2013_[130], Hoeymans 외, 2004_[154]).

²² 응답자가 PHQ-8의 처음 두 문항 중 한 문항에 ‘절반 이상’으로 응답하고 8개 항목 중 총 5개 이상의 항목이 ‘절반 이상’으로 보고되는 경우, 주요우울장애 위험군으로 간주된다(Eurostat, n.d._[74]).

²³ 준거점수 사용의 일관성이 부족하면 해석의 혼란을 초래할 수 있다. 예를 들어 2019년 미국 로스앤젤레스에 위치한 세 기관에서 노숙자 인구의 정신건강 상태 유병률 추정치를 발표했는데, 매우 상이한 결과가 나타났다. 로스앤젤레스 타임즈, 로스앤젤레스 노숙자 서비스 기관, 캘리포니아대학교 로스앤젤레스의 캘리포니아 정책 연구소에서 발표한 추정치는 각각 67%, 29%, 78%로 나타났다. 이 세 기관에서 발표한 추정치는 모두 동일한 데이터 세트에서 도출되었으며, 이러한 차이는 통계적 해석에서 비롯된 것으로 보인다(Smith 및 Oreskes, 2019_[157]).

²⁴ 다른 연구자들은 SWEMWBS 척도에 대해 고정 준거점수를 제시했으며 정신적 웰빙 수준 낮음(7.00점~19.98점), 정신적 웰빙 수준 중간(19.99점~29.30점), 정신적 웰빙 수준 높음(29.31점~35.00점)으로 구분된다. 이러한 범주는 덴마크 인구에 대한 이전 연구에서 도출되었다. 낮은 수준의 정신건강은 분포의 하위 15% 백분위수, 높은 수준의 정신건강은 분포의 상위 15% 백분위수에 해당한다(Santini 외, 2022_[148]). WEMWBS에 대한 고정 준거점수는 개발되지 않았다.

²⁵ PHQ-4와 PHQ-8 또는 PHQ-9에서 도출된 주요우울장애 위험 또는 우울 증상 위험에 대한 통계를 비교하는 경우 주의가 필요하다. PHQ에는 여러 채점 방식이 있어 채점 방식에 따라 유병률 추정치가 달라질 수 있기 때문이다. PHQ-4와 PHQ-8에 모두 나타나는 두 가지 개별 지표에서 우울증 위험을 계산하여 직접 비교 가능한 추정치를 산출할 수 있다. 다른 (과거) 건강 보고에서 우울증 위험 유병률을 추정하기 위해 PHQ-8의 전체 지표를 사용했더라도 해당 방식으로 종합 사회 조사 및 건강 조사 간의 측정값을 완전히 일치시킬 수 있다.

인구 정신건강 측정

우수한 정신건강은 사람들의 웰빙에 있어 매우 중요한 부분이며, 코로나19 팬데믹으로 인해 그 중요성이 다시금 주목받게 되었다. 그러나 지금까지 각국 정부에서는 보다 광범위한 인구 수준에서 정신건강을 효율적으로 모니터링하는 방법과 정신질환과 긍정적인 정신 상태를 모두 고려한 모니터링 방법을 중점으로 충분한 논의가 이루어지지 않았다. 본 보고서는 국가 통계청과 기타 데이터 생산기관에서 조사의 빈도를 높이고 국제적으로 일관성 및 통일성 있는 방식으로 조사를 수행하여 고품질의 인구 정신건강 결과 측정을 제공할 수 있도록 지원한다. 본 보고서는 OECD 국가의 기존 측정 관행을 문서화하고 사용 가능한 측정 도구의 장점 및 한계점을 논의하며 가정, 사회 및 건강에 관한 조사에서 채택할 우선순위 측정에 대한 권장사항을 제공한다. *인구 정신건강 측정*은 OECD의 웰빙 측정 연구의 맥락에서 정신건강 및 웰빙에 대한 평가의 일환으로 작성된 두 개의 보고서 중 첫 번째 보고서이다.

