

우울 증상과 수준에 따른 패시브 센싱의 패턴 조사[†]

정 승 민

고려대학교 안암병원 수련생

정 경 미[‡]

연세대학교 심리학과 교수

패시브 센싱은 스마트 기기를 통해 행동, 생리, 사회, 환경적 지표를 객관적이고 자동적으로 수집하는 방법으로, 최근 정신장애 평가 도구로서의 가능성이 탐색되고 있다. 그 중 우울장애에 대한 연구가 가장 활발한데, 우울장애는 증상과 수준에 따라 발현 양상이 상이할 수 있음에도 불구하고, 센서데이터를 이용한 연구에서는 이런 측면에 대한 고려가 제한적이다. 본 연구에서는 우울 증상별 관련이 높은 센서데이터를 규명하고, 우울 수준에 따라 센서데이터의 차이가 있는지를 확인하고자 하였다. 이를 위해 총 64명의 대학생 및 대학원생을 우울 수준에 따라 세 가지 집단(정상군, 경도 우울군, 고도 우울군)으로 분류하고, 총 30일간 우울장애의 9개 증상에 대한 자기보고 자료를 1일 4회, 반복적으로 수집하였다. 동시에, 스마트폰 및 스마트워치를 사용하여 총 8개의 센서데이터를 항시 수집하였고, 여기서 총 14개의 특성정보를 1일 4개의 값으로 추출하였다. 연구 결과, 각 증상별로 서로 다른 특성정보가 유의미한 관계를 가지는 것으로 나타났는데, DSM-5에 따른 9개의 증상 중, 2개 증상(흥미 저하, 피로감)은 모든 특성정보와 관련 있고, 5개 증상(우울한 기분, 수면 문제, 집중력 저하, 정신운동 문제, 자살사고)은 일부 특성정보와만 관련 있었지만, 나머지 2개 증상(식욕 문제, 무가치감)은 어떠한 특성정보와도 관련 없는 것으로 나타났다. 또한, 우울 수준에 따라 모든 특성정보 값에 유의미한 차이가 있는 것으로 나타났는데, 집단 차의 세부양상은 특성정보마다 달랐다. 활동성 및 이동성 지표는 정상군에 더 높게 나타났고, 휴대폰 잠금해제 기록과 평균 수면 시간과의 차이는 우울군에서 더 높게 나타났으며, 통화기록과 방문 장소의 개수는 경도 우울군에서 가장 높게 나타났다. 본 결과를 바탕으로, 연구의 임상적 함의와 한계점에 대해 논의하였다.

주요어: 패시브 센싱, 우울, 증상, 수준

[†] 본 연구는 2020년도 정부(과학기술통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구이며 (No.20200019650022002, 디지털 피노타이핑 기반 다차원적 우울 탐지 플랫폼 개발), 1저자의 석사 학위 청구 논문을 수정·정리한 것임

[‡] 교신저자(Corresponding author): 정경미, (03722) 서울특별시 서대문구 연세로 50, 연세대학교 심리학과 교수, Tel. 02-2123-2448, E-mail: kmchung@yonsei.ac.kr

디지털 기술의 급격한 발전과 함께 휴대용 스마트 기기의 보급이 증가함에 따라, 모바일 헬스(mobile Health; mHealth)의 일환으로 디지털 표현형(digital phenotyping)이라는 개념이 정신건강 관리 분야에 소개되었다. 디지털 표현형이란, 스마트폰과 같은 개인용 디지털 기기를 사용하여 개인으로부터 고유한 다양한 정보를 매 순간마다 수집하여 정량화하는 것을 뜻한다(Torous, Kiang, Lorme, & Onnela, 2016). 디지털 표현형은 전통적인 정신건강 평가도구만으로는 완전히 파악하기 어려운 행동, 생리, 환경, 사회적 지표에 대한 다차원적이고 대량의 정보를 제공할 수 있다는 점에서, 정신건강 측정을 위한 새로운 도구로서의 가능성이 탐색되고 있다(Insel, 2017). 현재까지 관련 연구는 초기 단계로, 임상적 타당성이나 개인 정보 보호 등 중요한 이슈들에 대한 해결이 필요하다(Mohr, Zhang, & Schueller, 2017). 그러나, 반복적인 연구를 통해 유용성이 축적된다면 연구 측면에서 실시간으로 대량의 행동 및 생리 정보 수집을 가능하게 하여, 차원적 진단체계 구축을 추진하는 Research Domain Criteria(RDoC; Insel et al., 2010) 뿐 아니라 평가와 측정 영역에서 새로운 프레임을 제시할 것이다(Torous, Onnela, & Keshavan, 2017). 또한, 임상 측면에서는 환자 개인의 상황을 고려하여 진단의 정확도 향상을 도모하는 정밀 의학(precision medicine)의 도구로서 기능할 것으로 기대된다(Insel, 2017).

디지털 표현형에서 수집되는 자료는 수집에 요구되는 개인의 노력에 따라 액티브 자료(active data)와 패시브 자료(passive data)로 구분된다(Onnela & Rauch, 2016). 전자는 생태순간평가(Ecological Momentary Assessment: EMA)와 같

이 본인의 경험에 대한 반복적인 자기보고 형태로, 후자는 모바일 및 웨어러블(wearable) 기기에 내장된 여러 센서(sensor)를 통해 자동적으로 수집된다. 이 중, 패시브 센싱(passive sensing)은 기존 도구와 달리 기기를 통해 측정된 객관적인 센서데이터에 근거하고, 자기보고로는 파악하는 데 한계가 있는 다양한 지표들을 자동적으로 수집하여 개인의 표현형을 구성한다는 점에서 특히 주목받고 있다(Huckvale, Venkatesh, & Christensen, 2019; Rohani, Faurholt-Jepsen, Kessing, & Bardram, 2018). 최근 10년간 급격히 활발해진 패시브 센싱 관련 연구는 우울장애(Ghandeharioun et al., 2017; Place et al., 2017; Yue et al., 2018), 양극성장애(Beiwinkel et al., 2016; Faurholt-Jepsen et al., 2019), 조현병(Shin et al., 2016; Tron, Resheff, Bazhmin, Peled, & Weinsall, 2017), PTSD(Balbin et al., 2017; McDonald, Sasangohar, Jatav, & Rao, 2019) 등 다양한 정신장애에 적극적으로 도입되고 있다.

우울장애는 높은 유병률, 공병률, 그리고 재발율로 인해 만성적이고 심각한 경과를 보이는 만큼, 관련 연구의 진행이 가장 활발하다(Liang, Zheng, & Zeng, 2019). 패시브 센싱에 기반하여 우울 측정을 시도한 선행연구는 최근 급속히 증가하고 있는데, 대다수의 연구는 EMA를 통해 수집한 액티브 자료를 준거로 센서데이터의 타당도를 확인했다는 측면에서 공통점을 지니지만, 수집한 센서데이터의 종류와 연구설계 및 분석방법 등에서 차이를 보인다. 가장 초기 연구들은 주로 단일 센서만을 사용하였다. Canzian 등(2015)은 46명의 성인을 대상으로 약 두 달간 GPS 정보를

수집한 결과, 개인의 활동성이 EMA를 통해 매일 측정된 우울증 선별도구(Patient Health Questionnaire: PHQ-8; Kroenke et al., 2009) 점수와 높은 부적 상관이 있음을 밝힌 한편, Mehrotra 등(2015)은 25명의 성인을 대상으로 30일간 휴대폰 사용 기록을 수집하여, 매일 PHQ-8로 측정된 우울 수준과 ‘휴대폰 알림에 대한 반응 시간’ 간 유의미한 정적 상관을 보고하였다.

후속 연구자들은 두 가지 이상의 센서데이터를 조합하여 우울을 연구하기 시작하였는데, Saeb 등은 총 3편의 연구를 통해 GPS 정보 및 휴대폰 사용 기록과 우울 간의 관계를 살폈다. 먼저, 40명의 성인을 대상으로 2주간 데이터를 수집하여 사전·사후에 측정된 PHQ-9 점수와 상관을 관찰한 결과, 우울 수준이 높을수록 더 적은 개수의 장소를 방문하고, 휴대폰을 더 많이 사용하는 경향이 있는 것을 확인하였다(Saeb et al., 2015). 이어서 48명의 대학생을 대상으로 10주간 GPS 정보를 수집하여 사전·사후의 PHQ-9 점수와 상관을 분석한 결과, 주중보다 주말에 두 정보 간 상관이 더 높음을 확인하였다(Saeb, Lattie, Schueller, Kording, & Mohr, 2016). 마지막으로, 208명의 성인을 대상으로 6주간 수집한 GPS 정보와 휴대폰 사용 기록을 조합하여 총 11가지의 ‘의미있는 장소(semantic location)’ 정보를 추출하고, 각 장소에서 보낸 시간과 총 3회(0, 3, 6주차)에 걸친 PHQ-9 점수와 상관을 분석한 결과, 각 주차별로 유의미한 관계를 보이는 장소가 다르다는 다

소 비일관적인 결과를 보고하였다(Saeb, Lattie, Kording, & Mohr, 2017).

최근의 연구들에서는 센서데이터 종류의 확장 뿐 아니라, 현행 정신과적 진단체계의 적용이 이루어지고 있다. 한 연구진은 DSM-5 장애에 대한 구조화된 임상적 면담(Structured Clinical Interview for DSM-5 Disorders: SCID-5; First et al., 2016)으로 분류된 정상/우울 집단을 구분하는 데 PHQ-9 점수와 센서데이터(활동 종류 및 GPS 데이터)를 순차적으로 투입한 결과, 센서데이터가 PHQ-9에 더해 29%의 추가적인 설명력이 있음을 밝혔다(Farhan, Yue et al., 2016). 또 다른 연구에서는 머신러닝 모델을 사용하여 피부전도도, 수면, 움직임, 휴대폰 사용 기록, 장소 등의 센서데이터를 통해 해밀턴 우울증 평가척도(Hamilton Depression Rating Scale: HDRS; Hamilton, 1967)로 측정된 우울을 예측하고자 시도하였다(Ghandeharioun et al., 2017). 이 때 5개의 서로 다른 머신러닝 모델과 그것을 모두 결합한 앙상블 모델을 포함하여 총 6개의 모델을 비교하였는데, ‘앙상블(ensemble)’ 모델(4.5 RMSE¹⁾이 나머지 5가지 머신러닝 모델(4.6~5.5 RMSE)에 비해 가장 낮은 오차율을 보였다.

이러한 일련의 연구들은 센서데이터가 전통적인 우울 자기보고 척도의 점수와 높은 관련성을 가짐을 보고하였다. 그러나 이 자료는 우울과 센서데이터의 관련성을 제한적으로 보여주는데, 연구된 센서데이터의 종류와 준거자료의 측정 빈도

1) 평균 제곱근 오차(Root Mean Square Error: RMSE)란, 모델 예측 값과 실제 값 간 차이의 제곱 평균을 구하고 그것을 제곱근하여 산출되는 지표로, 모델의 정확도를 평가하고 다른 모델과 비교하기 위한 기준으로 사용된다. 해당 연구에서는 데이터의 90%를 사용해 도출한 예측 값과 나머지 10%의 실제 값을 비교하였다. 이는 데이터의 척도에 따라 값이 달라지기 때문에 절대적인 기준은 없으며, 모델 간 상대적인 비교만 가능하다.

(예, 최대 1일 1회)에서의 제한 뿐 아니라, 복잡한 우울을 단일하게 개념화한 것 등이 문제로 지적된다(Mohr et al., 2017; Raballo, 2018). 그 중 우울의 단일한 개념화는 가장 많이 비판받고 있는데(Place et al., 2017), 일반적으로 우울은 증상과 그 심각도에 따라 발현 양상이 매우 상이하여 (Penninx, Milaneschi, Lamers, & Vogelzangs, 2013; Zimmerman, Balling, Chelminski, & Dalrymple, 2018) 다양한 증상의 조합으로 발현될 수 있는 다차원적인 장애로 개념화되기 때문이다 (Reinertsen & Clifford., 2018). 그러나, 현재 패시브 센싱을 이용한 대부분의 연구들은 자기보고 척도의 총점을 사용하는데, 이는 우울을 단일하게 개념화하는 것으로 세부적인 증상이나 우울 수준에 따른 차이에 대한 고려가 부족하다.

최근에는 이러한 비판을 고려하여 센서데이터와 우울의 관련성을 연구할 때, 우울을 총점으로만 측정하지 않고, 세부적인 증상 수준에서 조사하고 있다. Place 등(2017)은 정신질환 진단 및 통계 매뉴얼 5판(Diagnostic and Statistical Manual of Mental Disorders: DSM-5; American Psychiatric Association, 2013)에서 정의하는 우울의 진단기준에 포함되는 9가지 증상 중, 우울한 기분, 흥미 저하, 피로감이라는 세 가지 증상에 대한 정보를 포함하여 우울을 측정하였다. 이들은 우울의 필수 증상 2가지(우울한 기분, 흥미 저하) 중 1가지 이상을 경험한다고 대답한 73명의 성인을 대상으로 12주간 9가지 센서데이터(예, 신체활동, 장소, 휴대폰 사용 기록, 음성)를 수집한 후, 대상자들에게 SCID-5를 실시하여 각 증상의 유무를 판단하였다. 머신러닝 기법을 사용하여 여러 센서데이터들의 조합을 투입한 결과, 우울한 기분

은 음성 정보를 통해(정확도 .74), 흥미 저하는 문자 횡수와 총 이동 거리의 조합을 통해(정확도 .75), 피로감은 통화 횡수와 문자횡수의 조합을 통해(정확도 .57) 가장 정확하게 예측할 수 있는 것으로 나타났다.

Wang 등(2018)의 연구에서는 DSM-5 진단기준 중 앞선 세 가지 증상에 더하여 수면 변화, 집중력 저하를 포함하여 총 5가지 증상으로 우울을 개념화하였다. 83명의 대학생을 대상으로 겨울(56명) 및 봄 학기(27명)에 각각 9주 동안 활동 종류, 대화, 장소와 이동성, 수면 패턴, 심박수 등의 총 8가지 센서데이터를 수집하였으며, 우울의 각 증상에 대한 문항을 포함하는 PHQ-8를 사전·사후에, 그리고 PHQ-4을 주 1회 실시한 후, 일련의 분석을 진행하였다. 먼저, 센서데이터와 PHQ-8 점수와의 상관 분석을, 반복 측정된 PHQ-4 점수와의 일반화선형혼합모형을 진행한 결과, 우울의 각 증상마다 서로 다른 센서데이터가 관련 높은 것을 확인하였다. 예를 들어, 우울한 기분은 휴대폰 잠금해제 시간과 정적 관계, 흥미 저하는 정지 시간과는 정적 관계, 이동성과는 부적 관계를 보였다. 수면 변화는 잠금해제 시간과 정적 관계, 수면 시간과 부적관계를 보였고, 피로감은 기숙사에 있는 시간과 부적 관계, 집중력 저하는 잠금해제 시간과 정적 관계를 보였다. 또한, PHQ-8 상 10점을 기준으로 나눈 두 집단 간 센서데이터를 비교한 결과, 우울 집단에서 ‘공부 장소에서의 휴대폰 잠금해제 시간’이 유의미하게 더 높게, ‘대화 시간’이 유의미하게 더 낮게 나타났다. 마지막으로, 머신러닝 기법을 사용하여 사전·사후의 PHQ-8 점수를 가장 잘 설명하는 센서데이터를 선별하여 투입한 결과, 각각을 74%, 58%로 예측

함을 확인하였다.

또 다른 일련의 연구들은 우울의 수준을 고려하여 우울 집단을 정상 집단과 비교하였다. 앞서 언급된 Saeb 등의 연구(2015)에서는 PHQ-9 점수와 GPS 데이터 간 상관을 바탕으로 GPS 데이터로 우울 예측을 시도하였다. 그 결과, 86.5%의 정확도로 PHQ-9의 총점을 기준으로 나눈 정상과 우울 집단을 분류할 수 있었다. 이후에는 보다 많은 센서데이터를 사용하여 집단 분류를 시도하였는데, Wahle 등(2016)은 머신러닝 기법을 사용하여 신체활동, WiFi, GPS, 휴대폰 사용 기록 등의 센서데이터가 PHQ-9 총점에 따라 분류된 정상과 우울 집단을 약 60%의 정확도로 예측함을 보고하였다. Farhan 등(2016) 역시 머신러닝 기법을 사용하여, 조도, 휴대폰 사용 기록, 음성 정보들을 포함한 센서데이터의 패턴에 따라 대상자들을 세 가지 집단으로 분류하였고, 이 집단들이 각각 낮은, 중간, 높은 PHQ-9 점수를 보임을 보고하였다 (Farhan, Lu et al., 2016).

상기 연구들은 우울을 보다 세분화하여 살펴보았다는 점에서 새로운 시도이나, 우울과 센서데이터 간의 관련성을 밝히는 데에는 여전히 부족하다. 첫째, 현재 정신장애 진단의 기준으로 가장 통용되고 있는 DSM-5에서 정의하는 우울의 개념을 측정하기 위해선 DSM-5에 포함되는 9가지 증상 모듈을 고려할 필요가 있다. 둘째, 우울은 심각도에 따라 수준별로도 서로 다른 양상을 보인다는 점(Zimmerman et al., 2018)을 고려할 때, 선행 연구의 이분법적 집단 구분(정상/우울)에서 나아가 우울 수준의 세분화가 필요하다. 마지막으로, 자기보고 자료가 항상 수집되는 센서데이터의 준

거자료로서 보다 타당하게 기능하고 하루 내에서도 변동성을 보이는 ‘우울’이라는 구인의 특성을 더욱 타당하게 측정하기 위해서는 그 측정간격을 기존 연구들의 최소 간격인 1일 1회보다 더 빈번하게 하는 것이 권고된다.

따라서 본 연구에서는, 하루에 반복적으로 수집한 자기보고 자료의 각 문항 점수를 준거로 하여 우울 증상별 관련이 높은 센서데이터를 규명하고, 우울 수준에 따라 센서데이터의 차이가 있는지를 확인하고자 하였다.

방법

연구 대상

본 연구의 참가자는 서울시 소재 1개 대학교에 재학 중인, 우울을 경험하거나 경험하지 않는 대학생 및 대학원생으로, 해당 대학의 커뮤니티²⁾를 통한 모집문 공고를 통해 모집되었다. 연구 참가 신청 건수가 가용한 스마트워치(총 20개) 보유 대수보다 많을 것으로 예상되어, 연구 참가 시작 시기를 나누어 모집하였다. 참가자들에게는 실제 연구 목적과 동일하게 설명되었으며, 요구특성 효과를 방지하고자 집단간 처치의 차이가 없으며 참가자들의 특성에 따른 선별 가능성의 차이가 없다는 점을 함께 안내하였다.

연구에 포함된 최종 참가자는 다음과 같은 절차에 의해 선정되었다. 먼저 1단계에선 정상군과 우울군을 선정하기 위해 PHQ-9을 실시하였다. 이는 우울장애의 유병률을 고려할 때, 우울 비경험자의 참가 신청이 상대적으로 많을 것으로 예상

2) 세연넷 www.seiyon.net

되었기 때문에, PHQ-9의 0~4점에 해당하는 참가자를 정상군으로, 5~27점에 해당하는 참가자를 우울군으로 분류하였다(Park, Choi, Choi, Kim, & Hong, 2010). 2단계로 참가자의 우울 수준에 따른 세부 집단을 분류하기 위해 BDI-II를 사용하였다.

총 326명의 연구 참가 신청자를 대상으로 스크리닝 평가를 진행한 결과, 다음의 기준을 충족한 108명을 면담 대상자로 선정하였다: 1) 연구 참가에 자발적으로 동의한 경우, 2) 안드로이드 운영체제의 스마트폰을 소지한 자. 다음으로, 108명의 면담 대상자들에 대하여 SCID-5와 한국판 백 우울 척도 2판(Korean-Beck Depression Inventory-II: K-BDI-II)를 실시하였고, 아래의 기준에 따라 세 가지 우울 수준에 속하는 최종 참가자 76명을 선정하였다: 정상군($n=23$), 경도 우울군($n=27$), 고도 우울군($n=26$). 먼저, 정상군의 선별기준은 다음과 같았다: 1) PHQ-9 총점이 0~4점이며 2) K-BDI-II 총점이 0~9점이고 3) SCID-5 상 어떤 장애의 기준을 충족하지 않는 경우. 경도 우울군의 경우, 다음을 충족하는 대상자를 선정하였다: 1) PHQ-9 총점이 5~27점이며 2) K-BDI-II 총점이 10~23점이고 3) SCID-5 결과, 주요우울장애(부분 관해), 지속성 우울장애(순수한 기분저하 증후군 동반), 지속성 우울장애(간헐적인 주요우울삽화, 현재 삽화를 동반하지 않는 경우) 중 1개를 충족하는 경우. 마지막으로, 고도 우울군의 선별기준은 다음과 같았다: 1) PHQ-9 총점이 5~27점이며 2) K-BDI-II 총점이 24~63점이고 3) SCID-5 결과, 주요우울장애, 지속성 우울장애(간헐적인 주요우울삽화, 현재 삽화를 동반하는 경우), 지속성 우울장애(지속성 주요우울 삽화 동반) 중 1개를 충

족하는 경우. 경도 우울군과 고위험군의 경우, SCID-5에 따라 다음의 기준 중 적어도 1개를 충족할 시 연구 대상에서 제외되었다: 1) 현재 또는 과거에 정신병적 증상을 보인 경우, 2) 현재 또는 과거에 경조증 또는 조증 삽화를 보인 경우, 3) 현재 적응장애로 인해 우울 기분을 보고하는 경우.

최종 연구 참가자 중 휴대폰 교체 및 여행 등의 이유로 중도 포기한 4명과 스마트워치 장기 미착용, 불성실한 설문 응답 등의 이유로 20% 이상의 데이터가 수집되지 않은 8명을 제외하고, 총 64명(남 30명, 여 34명)이 최종 분석에 포함되었다. 최종 분석에 포함된 참가자 정보를 표 1에 제시하였다.

본 연구는 연세대학교 생명윤리위원회(Institutional Review Board)의 승인을 받은 후(승인번호: 7001988-202006-HR-606-12) 진행되었다.

표 1. 참가자 성별 및 나이

구분	참가자($N = 64$ 명)		
	정상군 ($n = 20$)	경도 우울군 ($n = 24$)	고도 우울군 ($n = 20$)
성별(남/여)	11/9	11/13	8/12
평균 연령(SD)	24.72세 (3.25)	24.60세 (4.94)	22.95세 (2.48)

측정 도구

면담 도구

DSM-5 장애에 대한 구조화된 임상적 면담 임상가용(Structured Clinical Interview for DSM-5 Disorders Clinician Version: SCID-5-CV). DSM-5. 진단기준에 근거하여 참

가자들의 현재 우울 수준을 진단하고, 기타 정신 장애 공병의 유무를 확인하기 위해 DSM-5 장애에 대한 구조화된 임상적 면담 임상가용(SCID-5-CV DSM-5® 장애에 대한 구조화된 임상적 면담, 오미영, 박용천, 오상우, 2017)을 사용하였다. SCID-5-CV는 다음의 DSM-5 주요 장애에 대해 진단기준에 근거한 기준 항목으로 구성되며, 각 항목별로 -(없음/역치 미만)와 +(역치)를 평정하였다: 우울장애, 양극성장애, 조현병 및 기타 스펙트럼 장애, 물질사용장애, 불안장애, 강박장애, 외상후 스트레스 장애, 성인 주의력결핍/과잉행동장애, 기타 현재 진단되는 장애들, 적응장애. 순차적인 흐름 과정에서 주어진 지침에 따라 더는 평정될 필요가 없거나 관련이 없는 항목은 건너뛰기를 하였다. 모든 장애에 대한 평가가 끝난 후, 각 장애별로 현재 및 과거의 충족 여부를 판단하여 SCID-5-CV 진단적 요약 채점지를 작성하였고, 그를 바탕으로 우울장애 관련 진단적 인상을 내리고 기타 공병 유무를 확인하였다.

자기보고식 척도

한국어판 우울증 선별도구(Patient Health Questionnaire 9: PHQ-9). 연구 참가 신청자들의 우울 수준을 측정하여 스크리닝하고, 연구 참가자들의 시간대별 우울 수준을 반복적으로 측정하기 위한 EMA로 환자 건강 질문지(Patient Health Questionnaire 9: PHQ-9; Kroenke, Spitzer, & Williams, 2001)를 한국어판으로 표준화한 도구(안제용, 서은란, 임경희, 신재현, 김정범, 2013)를 사용하였다. 한국어판 PHQ-9은 DSM-IV의 주요우울장애 증상 9가지를 각 문항에 반영하

고 있으며, 다른 우울증 선별 검사들에 비해 문항 개수가 적고, 민감도와 특이도가 높다는 장점이 있다. 총 9문항으로 구성되며, 지난 2주간의 상태에 대해 0점(전혀 없음)에서 3점(거의 매일)의 4점 리커트 척도로 평정된다. 최저 점수는 0점, 최고 점수는 27점으로, 점수가 높을수록 우울 정도가 심한 것을 의미한다. 한국어판 PHQ-9의 신뢰도 및 타당도 연구에서는 설문 총점에 따라 정상(0~4점), 경도(5~9점), 중등도(10~19점), 고도(20~27점)로 우울 수준을 분류하였다(Park, Choi, Choi, Kim, & Hong, 2010). 한국어판 타당화 연구(안제용 외, 2013)에서 보고된 내적 일치도는 .95였으며, 본 연구에서의 내적 일치도 계수(Cronbach's α)는 .89로 나타났다.

본 연구에서는 PHQ-9을 스크리닝과 EMA 도구로 사용하였는데, 스크리닝에서는 PHQ-9 원칙도를 그대로 사용하였다. EMA를 통한 반복적인 측정을 위해서, 기존 연구에서와 같이 원칙도 문항은 동일하나, 리커트 척도를 1점(전혀 아님)~5점(매우 해당함)으로 변형하여 사용하였다(Pratap et al., 2019). EMA의 경우, 각 참가자는 지난 설문 시점부터 현재까지 4시간 동안의 상태에 대해 평정하였다.

한국판 벡 우울척도 2판(Korean-Beck Depression Inventory-II: K-BDI-II). 연구 참가자들의 우울 수준을 측정하고자 벡 우울척도 2판(Beck Depression Inventory-II: BDI-II; Beck, Steer, & Brown, 1996)을 한국어로 번안한 도구(성형모 외, 2008)를 사용하였다. K-BDI-II는 우울 증상을 선별하거나 평가하기 위해 개발된 자기 보고식 척도로, 우울증의 정서적, 인지적, 동기

적, 신체적 증상 영역을 포함하는 총 21개 문항에 대해 0~3점의 4점 리커트 척도로 응답하도록 되어 있다. 최저 점수는 0점, 최고 점수는 63점으로 점수가 높을수록 우울 정도가 심함을 의미한다. Beck 등은(1996)은 총점에 따라 우울하지 않은 상태(0~9점), 가벼운 우울 상태(10~15점), 중한 우울 상태(16~23점), 심한 우울 상태(24~63점)으로 분류하였다. 한국어판 타당화 연구(성형모 외, 2008)에서의 내적 일치도는 .83이었으며, 본 연구에서의 내적 일치도 계수(Cronbach's α)는 .87였다.

센서데이터. 연구 참가자들의 행동, 생리, 사회, 환경적 정보를 수집하기 위해 연구 참가자 본인의 스마트폰 및 배부된 스마트워치³⁾에 내재되어 있는 센서를 사용하였으며, 각각에서 수집된 센서데이터는 YouNoOne 어플리케이션(Narziev et al., 2020⁴⁾)을 통해 본 연구진의 서버로 전송되어 저장되었다. 이 때, 센서데이터는 데이터의 양과 배터리 소모에 미치는 영향을 고려하여 각 EMA 실시 직전 30분 동안의 데이터만이 전송되었으며, 각 센서데이터의 특성에 따라 개별적으로 설정된 수집 단위에 기반하여 서버로 전송되었다. 서버에 전송되어 저장된 센서데이터는 데이터별로 특정 방법에 따라 사용자의 행동, 생리, 사회, 환경적 특성들을 나타내는 특성정보(feature)⁵⁾로 추출되었다. 대다수의 특성정보는 매 EMA 설문

의 간격에 맞추어 4시간 단위로 계산되어, 6~10시, 10~14시, 14~18시, 18~22시에 해당하는 시간당 1개의 값으로 추출되었으며, '평균 수면 시간과의 차이'만이 그 특성 상 하루에 1개 값으로 추출되었다. 특성 정보 중 시간 정보의 단위는 초(sec)였으며, 거리의 정보의 단위는 미터(meter)였다. 본 연구에서 사용한 센서와 그 수집 단위, 각 센서에서 추출된 특성 정보와 추출 방법을 표 2에 제시하였다.

이 때, 활동인식 API를 통해 수집 및 추출한 5가지 특성정보(정지 시간, 기울임 시간, 기립 시간, 이동 중 시간, 알 수 없는 움직임 시간)는 센서의 잦은 미작동으로 인해 데이터 수집률이 약 20%에 그쳐 분석에서 제외되었다. 또한, 어플리케이션 카테고리의 경우, 가장 대표적인 분류 기준으로 간주되는 구글 분류를 사용하고자 하였으나(표 3), 이 기준이 각 어플리케이션의 특성을 정확하게 반영하지 못한다고 판단하였고, 대안적인 분류기준 선정에 어려움이 있어 카테고리별 어플 사용 시간 역시 분석에서 제외하였다.

3) Samsung사의 Galaxy Gear S3의 Classic, Frontiers 모델과 Galaxy Active 모델을 사용하였다.

4) YouNoOne은 정보통신기획평가원(한국연구재단부설) 주관의 ICT혁신기술개발사업 아래 2019년 수행된 '우울 조기탐지 알고리즘 개발과 개입 ICT 서비스 플랫폼 구축' 과제의 일환으로 개발되었다. YouNoOne은 안드로이드 운영체제를 기반으로 하며, 스마트폰용과 스마트워치용 두 가지 버전으로 이루어져 있다. 참가자의 데이터는 개별적으로 생성된 계정별로 구분되어 저장되었다. 어플리케이션에 대한 자세한 사항은 Narziev et al(2020)에 설명되어 있다.

5) 특성정보(feature)란, 원 센서데이터를 연구 목적에 맞추어 가공하여 추출한 값이다.

표 2. 센서, 수집단위, 각 센서에서 추출한 특성 정보, 추출방법

센서	수집 단위	특성 정보	추출 방법
활동인식 API		정지 시간 기동임 시간 기침 시간 이동 중 시간 알 수 없는 움직임 시간	사용자의 움직임을 '정지, 기동임, 기침, 이동 중, 알 수 없는 활동'의 5가지 활동으로 인식함에 따라 각 활동의 시간을 추출
중요한 동작 센서		주요 동작 수	원 데이터값
계보기		걸음 수	원 데이터값
수/발신 통화 기록		통화 시간 통화 횟수	
휴대폰 잠금해제 기록	하루 24시간, 항시	휴대폰 잠금해제 시간 휴대폰 잠금해제 횟수	사진-기반(각 사진의 시간과 횟수)
		평균 수면 시간과의 차이	하루 중 연속적으로 가장 긴 '휴대폰 잠금 상태' 시간을 수면 시간으로 추출한 후, 각 사용자의 수면 시간과 평균 수면 시간(7시간) 간 차이의 절댓값을 계산
어플리케이션 사용시간 기록		사진 어플 사용시간 교육 어플 사용시간 여가 및 음악 어플 사용시간 음식 어플 사용시간 게임 어플 사용시간 건강 어플 사용시간 쇼핑 어플 사용시간 커뮤니케이션 어플 사용시간 여행 어플 사용시간	구글에서 각 어플별로 제공하는 카테고리(genre)에 따라 다음의 8가지 카테고리로 어플을 분류 후 각 카테고리별 어플 사용 시간을 추출: 1) 사진, 2) 교육, 3) 여가 및 음악, 4) 음식, 5) 건강, 6) 쇼핑, 7) 커뮤니케이션, 8) 여행. (각 카테고리의 예시 어플은 표 3에 제시되어 있음)
GPS	5분에 한 번씩, 1분 동안	이동 총 거리 장소 간 거리의 분산 방문 장소의 개수 두 장소 간 최장 거리 집에서의 최장 거리 이동 반경 ¹⁾	추적된 이동 경로에서 각 정보 추출 ²⁾
심박수 센서	5분에 한 번씩, 30초 동안	평균 심박수	30초간 측정된 심박수의 평균값

1) 이동 반경이란 특정 시간대 내에 방문한 장소들의 중심(centroid)에서 각 장소들 간 거리의 분산이다(Canzian et al., 2015).

2) 이동과 관련된 각 특성정보는 Canzian 외 동료들(2015)에서 사용한 공식을 그대로 사용하여 추출하였다.

표 3. 카테고리 별 예시 어플

카테고리	예시 어플
사진 (art & design)	카메라, 사진 편집 등
교육 (education)	외국어 공부, 서점, 대학 관련 어플, 공부 타이머 등
여가 및 음악 (entertainment & music)	음악, 라디오, 영화관, TV, 동영상 플레이어 등
음식 (food & drink)	식당 예약, 커피전문점 어플, 배달, 식당 평가 등
건강 (health & wellness)	운동, 수면, 다이어트, 뷰티 등
쇼핑 (shopping)	쇼핑몰, 은행, 투자증권 등
커뮤니케이션 (social & communication)	SNS, 채팅, 소개팅, 웹, 영상통화 등
여행 (travel)	지도, 교통수단 대여, 여행 가이드, 항공/숙박 어플 등

연구 절차

참가자 모집 및 스크리닝

온라인 링크를 통해 연구 참가 신청을 받음과 동시에 PHQ-9을 실시하였다. 설문 결과, 정상 및 우울 수준에 해당하는 참가자 수의 비율을 고르게 조정하여 면담 대상자를 선별하였다. 참가자들에게는 실제 연구 목적과 동일하게 설명되었으며, 요구특성 효과를 방지하고자 집단간 처치의 차이가 없으며 참가자들의 특성에 따른 선별 가능성의 차이가 없다는 점을 함께 안내하였다.

임상적 면담 및 사전평가

임상적 면담은 오프라인과 온라인⁶⁾의 두 방식으로 진행되었다. 오프라인의 경우, 연구진과 사전에 약속된 시간에 본 연구실을 방문한 참가자들을 대상으로 SCID-5-CV을 실시하였다. 온라인의 경우, 연구진과 사전에 약속된 시간에 화상 미팅 ZOOM⁷⁾에 접속한 참가자들을 대상으로 SCID-5-CV를 실시하였으며, 대면 검사와 최대한

동등한 환경을 갖추기 위해 피면담자와 면담자 모두 비디오를 켜 상태에서 진행하였다. 면담은 임상심리 전공 석사과정생 1인과 석사 후 연구원 1인에 의해 실시되었으며, 평가자는 모두 임상심리전문가 1인에 의해 훈련받은 후 실시에 임하였다. 두 평가자 모두 훈련가가 실시하는 면담을 관찰 후, 훈련가의 감독 하에 직접 면담을 실시하였으며, 면담이 종료된 후 필요에 따라 훈련가가 추가 정보를 탐색하였다. 평가자의 독립적인 실사가 가능할 때까지 훈련가가 면담을 감독하였다.

면담에 이어서, 참가자들은 서면(오프라인 면담의 경우)이나 온라인 링크(온라인 면담의 경우)를 통해 K-BDI-II에 응답하였다. 기존 연구들은 서면과 온라인 방식이 심리측정적인 속성 측면에서 차이가 없다고 보고한다(Holländare, Askerlund, Nieminen, & Engström, 2008). 연구진은 면담 및 설문 결과를 근거로, 앞서 언급된 참가자 선정 및 제외 기준에 따라 최종 참가자를 선별하였다. 이때, 국내 표본을 대상으로 K-BDI-II의 절단점을 도출한 연구들에서는 임상

6) 실험 초기에는 모든 면담이 오프라인으로 진행되었지만, COVID-19 사태로 인해 대면을 최소화하기 위해 IRB의 승인을 얻은 후 면담을 모두 온라인으로 전환하여 실시하였다.

7) ZOOM은 원격 회의 서비스를 제공하는 해외 플랫폼으로, 화상 회의, 채팅, 실시간 콘텐츠 공유 등의 기능을 지원한다.

군 선별을 위한 절단점만을 제시하였기에(성형모 외, 2008; 임선영 외, 2011), 세 집단에 대한 절단점으로는 국외 표본을 대상으로 한 연구결과를 차용하였다(Beck et al., 1996). 아울러 참가자들에게는 세 집단 중 어느 집단에 속해 있는지에 대해 은폐(blind)하였다.

오프라인으로 진행한 최종 참가자들은 면담 직후 스마트워치를 수령하여 어플리케이션을 설치하였고, 온라인으로 진행한 참가자들은 스마트워치 수령과 어플리케이션 설치를 위해 연구진과 시간 및 장소를 다시 한 번 정하여 연구실을 방문하였다. 설치 시, 먼저 스마트워치를 참가자들의 스마트폰과 동기화시킨 후, 연구진이 컴퓨터 프로그램⁸⁾을 사용하여 스마트워치용 YouNoOne 어플리케이션을 설치하였다. 동시에, 참가자의 스마트폰에 스마트폰용 APK⁹⁾ 파일을 설치하였고, 참가자들은 스마트폰용 YouNoOne에서 회원가입 후, 동일한 정보로 스마트워치용 YouNoOne에서도 로그인하였다. 마지막으로, 연구진이 어플리케이션 매뉴얼을 전달하며 간단한 사용방법 및 주의사항에 대해 설명하였다. 온라인 및 오프라인에 관계없이 면담에 소요된 시간은 총 40~50분이었고, 어플리케이션 설치에 소요된 시간은 약 30분이었다.

생태순간평가(EMA) 및 센서데이터 자료 수집

모든 연구 참가자들은 임상적 면담 및 사전 평가를 마친 다음 날부터 YouNoOne이 설치된 스마

트폰을 소지 및 스마트워치를 착용하고 총 30일간 일상생활을 하였다.

연구 참가자들은 자신의 스마트폰 내 설치된 YouNoOne을 통해, 하루에 4번, EMA로 실시한 설문문에 응답하였다. 설문은 정해진 시간(10시, 14시, 18시, 22시)에 팝업 알람과 함께 제공되었고, 알람 후 1시간 이내로 응답이 가능하였다. 설문 문항으로는 EMA 간격에 맞추어 변형된 PHQ-9을 사용하였다. 연구 참가자들의 EMA 참여율을 높이기 위해, 설문문에 응답한 비율에 따라 차별화된 참가 사례비를 지급하였다. 참여율과 상관없이 지급되는 기본 참가 사례비(20,000원)에 더해, EMA 1회 응답 시 250원을 책정하여 총 30일의 기간 동안 참가자의 응답률에 따라 최대 30,000원의 참가 사례비가 추가적으로 지급하였다.

센서데이터는 참가자가 스마트폰을 소지하거나 스마트워치를 착용하는 동안에 YouNoOne을 통해 자동적으로 수집되었다. 스마트폰 및 스마트워치는 하루 최소 8시간 이상 소지 및 착용할 것이 권장되었으며, 이를 미달하는 참가자에 한해서는 문자, 전화 등의 개별 연락을 통해 소지 및 착용을 장려하였다.

개인정보 보호를 위해, 수집되는 데이터를 서버로 전송시키는 과정에서 HTTP 프로토콜¹⁰⁾을 사용하여 데이터를 암호화(encrypt)하여 전송하고 클라우드에서 수신한 데이터를 복호화(decrypt)하여 이중암호화 처리하였다. 클라우드에 데이터를

8) 연구대상자의 스마트워치를 연구담당자의 컴퓨터에 연결시키는 데에는 Device Manager 프로그램이 사용되었으며, 스마트워치에 어플리케이션 파일을 설치하는 데에는 Visual Studio 2019가 사용되었다.

9) 구글(Google)사의 안드로이드 운영 체제에서 응용 프로그램과 미들웨어를 설치하고 배포할 수 있도록 된 파일 포맷이다.

10) 컴퓨터 네트워크 및 인터넷을 통한 보안 통신에 사용되는 프로토콜이다.

저장할 시에는 비크립트(bcrypt) 알고리즘¹¹⁾으로 데이터를 해싱(hashing)¹²⁾함으로써, 수집된 데이터를 사용자별로 구분할 수는 있지만 각 사용자가 누구인지는 식별할 수 없도록 처리하였다.

EMA 응답, 어플 활성화 상태, 데이터 업로드 현황 등을 포함한 모든 데이터 수집 현황은 관리자 프로그램¹³⁾을 통해 모니터링 및 관리되었으며, 데이터 수집 및 업로드에 문제가 발생한 참가자들에게는 개별 연락을 통해 필요한 절차를 안내하였다.

사후 평가

30일의 기간이 종료된 연구 참가자를 대상으로 오프라인과 온라인의 두 가지 형식으로 사후 평가를 진행하였다. 오프라인의 경우, 참가자는 30일 기간 종료 후 1주 이내에 본 연구실에 재방문하여 스마트워치를 반납하고 K-BDI-II에 응답하였다. 이후, 간단한 면담을 통해 연구 참가 기간 동안 긍정적이거나 부정적인 특별한 사건의 발생 여부와 관련 사실을 확인하였다. 온라인의 경우, 참가자는 30일 기간 종료 후 1주 이내에 온라인 링크를 통해 동일한 설문을 진행하였으며, 특별한 사건의 발생 여부 및 관련 사실 또한 링크의 주 관심 문항을 통해 응답하였다. 별도로, 스마트워치의 반납을 위해서는 연구실을 재방문하였다. 상기의 모든 절차를 완료한 참가자들에게는 EMA 참여율에 따라 해당하는 연구 참가비를 지급하였다.

분석 방법

본 연구의 통계 분석은 IBM SPSS(The Statistical Package for the Social Sciences) Windows Version 25.0을 이용하여 실시하였다.

먼저, 본 분석에 앞서 집단 간 동질성을 확인하기 위해 성별에 대해서는 카이제곱 검정을, 연령에 대해서는 일원배치 분산분석(one-way ANOVA)를 실시하였다. 아울러 EMA 응답 비율과 센서데이터 수집 비율에 따라 결과에 편향이 나타날 수 있는 점을 고려하여 집단간 EMA 응답 비율 및 센서데이터 수집 비율에 대해 일원배치 분산분석(one-way ANOVA)를 실시하였다.

본 분석 절차는 다음과 같다. 첫째, 자기보고로 수집된 총 9개의 우울 증상별 점수와 총 14개의 특성정보 간의 관계를 확인하기 위해 다음의 분석을 진행하였다. 먼저, 경도 우울군과 고도 우울군에 해당하는 44명을 대상으로 PHQ-9 각 문항 점수의 정규성 검정을 진행하여, 모든 문항에서 정규성 가정의 충족을 확인하였다. 이에, 해당 자료를 대상으로 반복측정 자료의 분석에 이용되는 선형혼합모형(linear mixed model: LMM)을 진행하였다. 정상군은 집단의 특성 상, PHQ-9에 대한 응답이 대부분 1점으로, 그 범위가 매우 제한되어 분석에서 제외되었다. 아울러 분석을 진행하기에 앞서, 자기보고 점수는 1~5점 범위인 반면, 각 특성정보의 값은 데이터의 종류에 따라 그 단위가 0~10,000까지 넓기에, 이 차이를 조정하기 위해 특성정보 값의 단위를 모두 2,000으로 나누

11) 해싱 알고리즘 중 하나로, 비밀번호 암호화에 특화되어 있는 기법이다(Provos & Mazieres, 1999).

12) 특정 키 값에 연산을 적용하여 키 값이 있는 기억 장소의 주소로 바로 접근하는 방법으로, 암호화에 사용된다.

13) 관리자 프로그램은 EasyTrack 프로그램을 통해 YouNoOne 어플리케이션과 연동이 되도록 제작되었으며, 참가자별로 EMA 응답, 마지막 YouNoOne 활성화 시간, 마지막 데이터 업로드 시간 등을 열람할 수 있다.

어 0~5점 사이로 변환하였다. 종속변인과 독립변인을 투입함에 있어, 하루에 추출된 특성정보의 개수에 따라 ‘평균 수면 시간과의 차이’와 나머지 13개의 특성정보가 다르게 분석되었다. ‘평균 수면 시간과의 차이’는 하루에 1개의 값만 추출되었기에, 이에 맞추어 종속변인으로는 ‘과다수면 또는 불면’ 문항 점수의 하루 평균값이, 독립변인의 고정효과(fixed effect)로는 ‘평균 수면 시간과의 차이’가 투입되었다. 나머지 13개의 특성정보들은 하루에 4개의 값이 추출되었기에, 종속변인은 10시, 14시, 18시, 22시에 수집된 PHQ-9의 각 문항 점수였다. 독립변인의 고정효과로는 각 PHQ-9가 실시되기 직전의 4시간(예, 6~10시, 10~14시, 14~18시, 18~22시) 당 1개 값으로 추출된 각 특성정보를 중심화(centering)한 값을 투입하였다. 센서데이터는 24시간 항시 수집되었지만, 준거자료의 간격에 맞추어 총 16시간의 데이터만이 분석에 사용되었다.

둘째, 우울 수준별 세 집단에 따라 각 센서데이터 값에서 집단 차이가 있는지 확인하기 위해 다음의 분석을 진행하였다. 먼저, 전체 참가자 64명을 대상으로 PHQ-9 각 문항 점수의 정규성 검정을 한 결과, 모든 문항에서 정규성 가정을 충족시키지 않음을 확인하였다. 이에, 해당 자료를 대상으로 일반화선형혼합모형(generalized linear mixed model: GLMM)을 진행하였다. 종속변인은 각 PHQ-9가 실시되기 직전 4시간의 특성정보를 중심화한 값이었으며, 독립변인은 우울 수준별 집단 정보였다. 수면 데이터는 앞선 분석과 마찬가지로 1일 1개의 값만 사용되었다.

결 과

집단 간 사전 동질성 검정

우울 수준에 따른 세 집단이 성별 및 연령에서 동질적인지 확인하기 위해, 사전 동질성 검정을 실시하였다. 먼저, 성별에 대하여 카이제곱 검정을 실시한 결과, 조건 간 성비에 유의미한 차이가 없는 것으로 나타났다, $\chi^2(3) = 10.60, p > .05$. 다음으로, 연령에 대해 일원배치 분산분석(one-way ANOVA)을 실시한 결과, 집단 간 연령($F(2, 58) = .88, p > .05$)에 유의미한 차이가 없는 것으로 나타났다. 마지막으로, EMA 응답 비율 및 센서데이터 수집 비율에 대해 일원배치 분산분석(one-way ANOVA)을 실시한 결과, 집단 간 EMA 응답 비율($F(2, 58) = .64, p > .05$) 및 센서데이터 수집 비율($F(2, 58) = .79, p > .05$)에 유의미한 차이가 없는 것으로 나타나 집단 간 사전 동질성이 확인되었다.

우울 증상별 관련 특성정보

경도 우울군 및 고도 우울군을 대상으로, 자기보고(PHQ-9)로 수집된 우울 증상별 점수와 각 특성정보 간의 관계를 확인하기 위해 회귀분석인 선형혼합모형을 진행하였다. 총 9개의 증상과 14개의 특성정보 간 관계를 분석한 결과, 각 증상별로 서로 다른 특성정보가 유의미한 관계를 가지는 것으로 나타났다(표 4). 문항 1(일상생활에 대한 흥미 저하)과 문항 4(피로감 또는 활력 저하)에 대해서는 모든 특성정보가 유의미한 관계를 보인 반면, $p < .01$, 문항 5(식욕 저하 또는 증가)와 문항 6(죄책감 또는 무가치감)에 대해서는 어떠한 특성정보도 유의미한 관계를 보이지 않았다,

표 4. 우울 증상별 특성정보의 비표준화된 회귀계수

구분	증상 1	증상 2	증상 3	증상 4	증상 5	증상 6	증상 7	증상 8
휴대폰 잠금해제 시간	-.27**	-.11		-.15**	-.03	-.01	.22**	-.09
휴대폰 잠금해제 횟수	.21**	-.02		.13**	-.03	.01	.11*	.05
통화 시간	.31**	.04		.22**	-.01	.02	.12*	.33**
통화 횟수	.43**	.01		-.22**	.02	-.09	.02	.17*
주요 동작 수	.15**	-.26**		-.11**	.01	.05	.29**	.01
걸음 수	-.21**	-.04		.31**	.04	-.04	.35**	-.03
이동 총 거리	-.21**	-.16**		.41**	.01	.04	.22*	.02
장소 간 거리의 분산	-.33**	-.13**		.32**	-.07	-.17	-.17**	-.03
이동 반경	-.22**	-.07		-.14**	.05	-.01	-.03	-.25**
방문 장소의 개수	-.14**	-.32**		.16**	.02	-.03	-.00	.01
두 장소 간 최장 거리	-.31**	-.21**		-.43**	-.00	.11	.07	-.17
집에서의 최장 거리	-.41**	-.11**		-.21**	.07	-.01	-.11*	-.01
평균 심박수	-.37**	-.21**		-.86**	.03	.09	.03	-.31
평균 수면 시간과의 차이			.12**					

* $p < .05$, ** $p < .01$.

표 5. 특성정보 별 집단의 주효과 및 사후분석 결과

특성정보	F	M [95% CI]			집단 비교 ^a
		정상군(A)	경도 우울군(B)	고도 우울군(C)	
휴대폰 잠금해제 시간	105.02**	1.19 [1.14, 1.23]	1.64 [1.59, 1.68]	1.26 [1.21, 1.31]	B > C > A
휴대폰 잠금해제 횟수	5.358**	3.33 [2.82, 3.85]	3.96 [3.46, 4.47]	4.50 [4.03, 4.98]	C > B > A
통화 시간	10.40**	1.70 [1.60, 1.80]	1.96 [1.88, 2.07]	1.71 [1.62, 1.81]	B > A, C
통화 횟수	3.75*	2.55 [2.52, 2.58]	2.60 [2.56, 2.63]	2.53 [2.50, 2.57]	B > A, C
주요 동작 수	10.04**	2.49 [2.26, 2.71]	2.66 [2.44, 2.89]	1.95 [1.71, 2.18]	A, B > C
결음 수	13.25**	1.29 [1.27, 1.30]	1.25 [1.23, 1.26]	1.23 [1.21, 1.24]	A > B > C
평균 수면 시간과의 차이	14.67**	1.11 [1.10, 1.12]	1.48 [1.46, 1.50]	1.73 [1.72, 1.75]	B, C > A
이동 총 거리	14.81**	2.37 [2.33, 2.41]	2.35 [2.31, 2.39]	2.23 [2.19, 2.27]	A, B > C
장소 간 거리의 분산	54.22**	2.33 [2.30, 2.37]	2.15 [2.11, 2.18]	2.08 [2.05, 2.01]	A > B > C
이동반경	7.49**	2.17 [2.16, 2.18]	2.16 [2.15, 2.17]	2.15 [2.15, 2.16]	A > B, C
방문 장소의 개수	6.69**	2.01 [1.89, 2.13]	2.34 [2.21, 2.47]	2.12 [1.98, 2.25]	B > A, C
장소 간 최장 거리	10.29**	1.04 [1.04, 1.04]	1.04 [1.04, 1.04]	1.03 [1.03, 1.03]	A, B > C
집에서의 최장 거리	89.90**	1.26 [1.23, 1.28]	1.07 [1.05, 1.09]	1.13 [1.11, 1.16]	A > C > B
평균 심박수	64.31**	2.03 [2.02, 2.05]	2.00 [1.99, 2.01]	1.98 [1.97, 1.99]	A > B, C

a. 집단 간 유의미한 차이를 부등호(<, >)로 표시함

* $p < .05$, ** $p < .01$.

$p > .05$. 문항 3(과다수면 또는 불면)은 ‘평균 수면 시간과의 차이’와 유의미한 관계를 보였으며, $p < .01$, 문항 2(우울한 기분), 7(집중력 저하), 8(정신운동 초조 및 지연), 9(자살사고 및 시도)는 특성정보 중 일부와만 유의미한 관계를 보였다, $p < .05$.

우울 수준에 따른 특성정보 차이

우울 수준에 따른 세 집단별로, 14개의 특성정보에서 차이가 있는지 확인하기 위해 회귀분석인 일반화선형혼합모형을 진행하였다. 분석 결과, 모든 특성정보에서 집단의 주효과가 유의미하였으며, $p < .05$, 사후분석 결과, 특성정보별로 집단 간 차이의 세부 양상이 달랐다(표 5). 먼저, 걸음 수, 장소 간 거리의 분산, 이동 반경, 평균 심박수는 우울군의 두 집단이 정상군에 비해 유의미하게 낮았으며, 주요 동작수, 이동 총 거리, 장소 간 최장 거리는 고도 우울군이 정상군 및 경도 우울군에 비해 유의미하게 낮았다. 반대로, 휴대폰 잠금 해제 시간 및 횟수와 평균 수면 시간과의 차이는 우울군이 정상군에 비해 유의미하게 높았다. 통화 시간 및 횟수와 방문 장소의 개수는 경도 우울군이 다른 두 집단에 비해 유의미하게 높았다.

논 의

본 연구에서는 자기보고로 측정된 각 우울 증상에 대해 유의한 관련을 보이는 센서데이터를 규명하고, 우울 수준에 따라 센서데이터 값에 차이가 있는지 살펴보았다. 이를 위해 연구 참여에 동의한 대학생 및 대학원생을 대상으로 PHQ-9, SCID-5-CV, BDI-II의 자료를 수집하고, 미리 정

한 3가지 집단(정상군, 경도 우울군, 고도 우울군)의 기준을 충족하는 64명을 선발하였다. 이들을 대상으로 총 30일간 1일 4회 EMA로 PHQ-9를 실시하여 9가지 우울 증상에 대한 자기보고 점수를 반복적으로 수집하였다. 동시에, 스마트폰 및 스마트워치를 사용하여 총 8개의 센서데이터를 항상 수집하였고, 그로부터 총 14개의 특성정보를 EMA 간격에 따라 1일 4개의 값으로 추출하였다.

이 자료를 사용해 우울 관련 자기보고 점수와 특성정보 간의 관계가 증상과 수준에 따라 어떻게 차이가 나는지 조사하였다. 연구 결과, 설명 변량이 작기는 하나 각 증상별로 서로 다른 특성정보가 유의미한 관계를 가지는 것으로 나타났는데, DSM-5에 따른 9개의 증상 중, 2개 증상(흥미 저하, 피로감)은 모든 특성정보와 관련 있고, 5개 증상(우울한 기분, 수면 문제, 집중력 저하, 정신운동 문제, 자살사고)은 일부 특성정보와만 관련 있었지만, 나머지 2개 증상(식욕 문제, 무가치감)은 어떠한 특성정보와도 관련 없는 것으로 나타났다. 또한, 우울 수준에 따라 모든 특성정보 값에 유의미한 차이가 있는 것으로 나타났는데, 집단 차의 세부양상은 특성정보마다 달랐다. 활동성 및 이동성 지표는 정상군에서 더 높게 나타났고, 휴대폰 잠금해제 기록과 평균 수면 시간과의 차이는 우울군에서 더 높게 나타났으며, 통화기록과 방문 장소의 개수는 경도 우울군에서 가장 높게 나타났다. 본 연구의 의의 및 함의는 다음과 같다. 첫째, 우울의 증상에 따라 서로 다른 센서데이터가 유의한 관련이 있음을 확인하였다. 우울장애의 각 증상들은 매우 이질적인 특징을 가지며(Otte et al., 2016), 동일한 진단 하에서도 개인이 보이는 증상의 조합이 매우 상이하다는 점에서(Fried &

Nesse, 2015), 기존 연구에서 주로 사용하였던 총점에 기반한 평가보다는 각 증상을 독립적으로 측정하는 시도가 제안되었다(Place et al., 2017; Wang et al., 2018). 이를 고려하여 본 연구에서는 기존 연구와 마찬가지로 PHQ-9을 사용하였으나, 총점을 사용하는 대신 DSM-5 진단기준 상 우울 장애의 9가지 증상에 각각 해당하는 9개 문항의 개별 점수를 사용하였다. 증상 개수가 일부로 한정된 기존 연구(Place et al., 2017; Wang et al., 2018)와는 달리, 본 연구에서는 9개 증상을 모두 조사하여 증상별로 서로 다른 센서데이터가 유의한 관계가 있음을 확인하였다는 점에서 기존 연구의 확장으로 볼 수 있다. 특히, 증상별로 센서데이터와의 관계를 조사한다는 동일한 접근을 취한 선행연구와 비교했을 때(Wang et al., 2018), 흥미저하와 피로감이 이동성과 부적 관계를, 집중력저하가 휴대폰 잠금해제 시간과 정적 관계를 보인다는 결과를 일관되게 얻었다.

결과를 좀 더 자세하게 들여다보면, 9가지 증상은 유의한 관계를 보인 특성정보의 개수에 따라 세 가지(모든 특성정보, 일부 특성정보, 없음)로 분류되었다. 이러한 결과는 본 연구에서 포함한 센서데이터의 종류와 관련 있는 것으로 보인다. 센서데이터는 그 특성에 따라 행동, 생리, 사회, 환경 등 여러 지표로 구분되는데(Rohani et al., 2018), 본 연구에서 사용한 14개의 특성정보는 다음과 같이 구분할 수 있다. 기기 사용 지표(휴대폰 잠금해제 시간 및 횟수), 사회성 지표(통화 기록 및 횟수), 활동성 지표(주요 동작 수, 걸음 수, 이동 총 거리), 이동성 지표(이동반경, 방문 장소의 개수, 장소 간 거리의 분산, 두 장소 간 최장 거리, 집에서의 최장 거리), 생리적 지표(심박수),

그리고 수면 지표(평균 수면 시간과의 차이). 이는 지표의 다양성 측면에서는 양호하지만, 다른 지표에 비해 활동성 및 이동성 지표에 해당하는 특성정보의 개수가 많아, 그 비중은 다소 편중되었다고 볼 수 있다. 이에, 각 증상별로 관련 있는 지표가 본 연구에 포함된 비중에 따라 결과가 다르게 나온 것으로 추정된다.

먼저, ‘일상생활에 대한 흥미저하’와 ‘피로감 또는 활력 저하’의 2개 증상은 모든 특성정보와 유의한 관계를 가지는 것으로 나타났다. 특히, 흥미저하와 피로감이 대부분의 활동성 및 이동성 지표와 부적 관계를 보였는데, 이는 흥미가 저하되고 높은 피로감을 경험할수록 행동의 강도와 범위가 축소되어 있음을 시사한다. 구체적으로, 총 걸음 수와 이동한 총 거리가 적고, 방문한 장소의 개수가 적으며, 그 장소들이 서로 넓게 분포되어 있지 않을수록 흥미 저하와 피로감의 수준이 심해졌다. 개인의 행동과 우울 증상 간 관계를 살핀 선행연구에서, 피로감과 이동성의 제한이 관련 있다고 보고된 것(Garg, Bush, & Gappmaier, 2016)과 신체활동 증진 개입이 흥미 저하를 포함한 우울 증상의 개선에 도움이 된다고 보고된 것(Dinas, Koutedakis, & Flouris, 2011)은 본 연구 결과를 지지한다. 한편, 두 증상 모두 휴대폰 잠금해제 시간과는 부적, 횡수와는 정적 관계를 보였는데, 동일한 기록에서 서로 다른 결과가 나타난 것은 두 특성정보가 개인의 행동에 대해 서로 다른 의미를 지니기 때문으로 보인다. 휴대폰 잠금해제 시간은 개인이 휴대폰을 사용한 시간으로, 일반적으로 개인의 여가 시간을 나타내는 반면, 그 횡수는 개인이 업무에서 벗어나 휴대폰으로 주의를 돌린 행동으로 간주할 수 있다. 이는 개인

이 흥미저하와 피로감을 높게 경험할수록, 주의집중하는 데 어려움을 겪고 실질적인 여가활동에 시간을 투자하는 데 실패함을 의미한다고 볼 수 있다.

다음으로, 5가지 증상(우울한 기분, 수면 문제, 집중력 저하, 정신운동 초조 및 지연, 자살사고 및 시도)은 일부 특성정보와만 유의한 관련성을 보였다. 먼저, 이동성이 낮고 평균 심박수가 낮을수록 개인이 우울한 기분을 더 많이 경험하는 것으로 나타났는데, 이는 선행연구와도 일치하는 결과이다(Canzian et al., 2015; Kemp et al., 2010). 개인의 이동성은 신체적 활동과도 관련 있는데, 움직이는 범위가 작고 활동이 적을수록 우울한 기분을 더 많이 느끼는 것으로 보인다(Peluso & Andrade, 2005). 다음으로, ‘과다수면 또는 불면’ 증상이 ‘평균 수면 시간과의 차이’와 유의한 정적 관계를 보여, 센서데이터를 통해 수면 불균형 문제를 유의하게 예측할 수 있음을 알 수 있었다. 해당 특성정보는 일차적으로 측정된 참가자 개인의 수면 시간과 평균 수면 시간으로 알려진 7시간과의 차이값으로 계산되었는데, ‘과다수면 또는 불면’ 증상 역시 수면 시간상의 문제를 다루고 있기에 두 변인 간 관계가 유의하게 나타난 것은 새롭지 않다. 최근에는 패시브 센싱을 통해 수면 시간 뿐 아니라 수면과 관련된 다양한 측면들(예, 깨시시간, 지연시간, 수면의 질 등)에 대한 측정이 시도되고 있는데(Kahveci, Alemdar, & Ersoy, 2015; Sano, Rahman, Zhang, Ganesan, & Choudhury, 2020), 이를 통해 보다 정확한 수면 측정이 가능해질 것으로 예상된다.

또한, 휴대폰 잠금해제를 많이 하고 활동성이 높을수록 집중력이 더 많이 저하되는 것으로 나

타났다. 높은 휴대폰 잠금해제 기록은 개인이 업무에서 자주 벗어날 때 나타나는 행동으로 간주되어 집중력 저하의 지표가 되는 것으로 추정되며(Wang et al., 2018), 높은 활동 수준은 높은 신체적 소모와 연결되어 집중력 저하를 야기하는 것으로 보인다. 한편, ‘정신운동 초조 및 지연’은 통화 기록 및 이동반경과, ‘자살사고 및 시도’는 다수의 데이터(예, 휴대폰 잠금해제 기록, 통화기록, 이동성 지표)와 유의미한 관련성이 나타났는데, 두 증상에 대해 조사한 선행연구는 매우 제한적이어서 관련성에 대한 근거를 찾아보기 어렵다. 일부 결과에 대해 추측해보자면, ‘자살사고 및 시도’가 휴대폰 잠금해제 횟수와는 정적관계를, 이동성 지표와는 부적 관계를 보인 것은 자살사고의 특성과 관련지어 생각해 볼 수 있다. 죽음에 대한 생각이 일상생활에서 반복적이고 침습적으로 나타날 경우, 개인은 학업 및 직업 장면에서 필요한 일에 집중하기 어려울 가능성이 있으며, 삶에 대한 낮은 동기 수준은 생활에 필요한 움직임을 제한하여 낮은 이동성으로 이어질 수 있다. 그러나, 두 증상에 대해서는 추가 연구를 통해 유의한 관계를 가지는 센서데이터에 대한 반복검증이 필요하다.

반면, ‘식욕 저하 또는 증가’와 ‘무가치감 또는 죄책감’은 어떠한 특성정보와도 유의한 관계를 가지지 않는 것으로 나타났는데, 이는 본 연구에 포함된 센서데이터와 관련되어 있는 것으로 생각해 볼 수 있다. 먼저, 식욕 문제의 평가를 위해서는 사용자의 실제 섭취 양상을 측정하는 것이 일반적인데, 실제로 일부 연구에서는 사용자가 사진이나 글로 남긴 음식 기록으로부터 섭취 빈도, 칼로리 등의 정보를 얻고 있다(Biel, Martin, Labbe, &

Gatica-Perez, 2018; Meegahapola, Ruiz-Correa, & Gatica-Perez, 2020). 본 연구에서는 해당 정보의 타당도나 정확도에 대한 이슈로 인해 이를 측정에서 제외하였는데, 추후 정확도 측면에서 기술적 보완이 이루어지게 되면 패시브 센싱을 통한 식욕 문제의 타당한 측정이 가능할 것으로 예상된다. 다음으로, ‘무가치감 또는 죄책감’은 우울장애의 9가지 증상 중, 인지적 요소를 가장 많이 포함하고 있는 증상으로 간주된다(Otte et al., 2016). 본 연구에서 수집한 센서데이터는 주로 행동적 특성과 관련된 지표(예, 휴대폰 잠금해제 기록, 활동성 및 이동성 지표)였는데, 이러한 센서데이터의 특성상 인지적 증상을 충분히 반영하지 못했을 가능성이 있을 것으로 추정된다. 인지는 행동과 비교해 내면화된 특성을 지녔기에, 개인의 내면을 보다 직접적으로 반영하는 데이터를 사용하는 것이 적절할 것으로 보인다. 최근, 음성 정보와 얼굴 표정에 대한 분석이 단순한 행동적 지표들과 비교해 개인의 내면에 대한 풍부한 정보를 제공한다는 점에서 우울 측정의 강력한 지표로 주목받고 있는데(Williamson et al., 2016; Place et al., 2017), 인지의 측정에도 두 데이터가 유용하게 사용될 수 있을 것이다.

둘째, 센서데이터로 우울을 측정할 때, 우울의 수준을 세분화하여 관찰할 필요성이 시사되었다. 정상군에 더해 우울을 두 수준으로 세분화하여 세 집단 간 차이를 관찰한 결과, 모든 특성정보에서 집단에 따른 차이가 유의미하였다. 특성정보별로 그 세부양상은 달랐는데, 대부분의 활동성 및 이동성 지표는 고도 우울군으로 갈수록 값이 낮아진 반면, 휴대폰 잠금해제 기록과 평균 수면 시간과의 차이는 값이 높아졌다. 일부 데이터(예, 방

문 장소의 개수, 통화 기록)은 경도 우울군에서 그 값이 가장 높게 나타났다.

먼저, 방문 장소의 개수를 제외한 모든 활동성 및 이동성 지표에서 경도 우울군과 고도 우울군이 정상군보다 유의하게 낮게 나타나, 개인이 우울할수록 신체활동과 이동성이 떨어짐을 관찰할 수 있었다. 우울한 사람들은 의욕이 없고 동기 수준이 현저히 저하되어 있는 경우가 많은데(Smith, 2013; Williamson & Shaffer, 2002), 이로 인해 신체적으로 활동하고 다른 장소로 이동하는 것에 대한 동기 역시 저하되고 해당 행동이 감소되었을 것으로 보인다. 또한, 작은 노력에도 큰 피로감을 경험하는 경향이 있어, 활동과 이동을 위한 에너지 자체가 부족하여 행동이 적게 나타났을 수 있다. 이러한 결과는 특히, 자기보고나 패시브 센싱을 통해 우울과 신체활동 및 이동성 간 부적 관계를 보고한 기존 연구결과와 일치하며(Canzian et al., 2015; Pickett, Yardley, & Kendrick, 2012; Vallée, Cadot, Roustit, Parizot, & Chauvin, 2011), 우울 관련 정보로서 활동성과 이동성 지표가 유용함을 시사한다.

반면, 휴대폰 잠금해제 기록과 평균 수면 시간과의 차이는 정상군에 비해 우울군에서 유의하게 높게 나타났는데, 우울할수록 휴대폰에 소비하는 시간이 많아지고, 수면 불균형 문제를 경험하는 것으로 보인다. 먼저, 다수의 우울한 사람들은 집중력 저하를 호소하는데, 이는 현재 하고 있는 일이나 공부에서 벗어나 다른 활동을 하는 경우를 말한다. 실제로, 기존 연구들에서 지나친 휴대폰 사용과 우울 간의 정적 관계가 일관적으로 보고된다(Coyne, Stockdale, & Summers, 2019; Tamura, Nishida, Tsuji, & Sakakibara, 2017). 또

한, 우울증의 증상 중 하나가 ‘과다수면 또는 불면’인 만큼, 우울을 경험하는 개인은 평균적인 시간의 수면을 제대로 취하지 못할 가능성이 크다. 이로 인해, 우울 수준이 높을수록 ‘평균 수면 시간과의 차이’ 특성정보 값이 크게 나타났을 것이다.

한편, 우울 수준과 선형적인 관계를 보인 대부분의 특성정보와 달리, ‘통화 기록’과 ‘방문 장소의 개수’는 정도 우울군에서 그 값이 가장 높게 나타났다. 우울 수준에 따라 표현형이 달라질 수 있는데, 상기 결과는 우울이 심각하지 않을 경우 일시적으로 사회적 관계를 더 추구하거나 활동성이 증가하는 현상을 반영하는 것일 수 있다. 그러나 이와 관련해 반복적인 연구를 통한 검증이 필요하다. 아울러 이러한 결과는 센서데이터를 통해 우울을 측정할 때, 데이터의 값이 높을수록 우울 수준이 높거나 낮다는 등의 일방향적인 해석으로 일괄하는 것에 주의를 요함을 시사한다. 센서데이터 개별의 집단 차 양상을 활용함으로써 데이터의 값으로부터 정확한 우울 수준을 도출해내는 것이 필요하다.

본 연구에서 우울을 2개의 집단으로 분류한 것은 제한된 참가자 수를 고려한 임의적 결정으로, 추후 연구에서는 우울을 더욱 세분화할 필요가 있을 수 있다. 실제 PHQ-9의 경우, 정상군을 포함해 최대 4개의 하위집단 분류가 가능하다(Park et al., 2010). 본 연구에서 2개의 우울 집단에서 수집한 센서데이터의 특성에 차이가 있는 것을 감안할 때, 집단을 보다 세분화하여 재탐색한다면 수준에 따른 센서데이터의 차이에 대한 유의미한 추가 정보를 얻을 수 있을 것으로 예상된다.

셋째, 본 연구에서 반복적인 자기보고 자료의 수집을 통해 준거자료의 타당도를 향상시켰다는

점은 주목할 만하다. 본 연구에서는 EMA 기법을 사용하여 PHQ-9 응답을 1일 4회 수집함으로써 그 측정간격을 최소화하였는데, 이는 준거자료의 타당도와 관련하여 두 가지 측면에서 이점을 지닌다. 먼저, 설문 실시 시간이 아침, 낮, 저녁, 밤 시간대(10, 14, 18, 22시)에 골고루 걸쳐있어 보다 다양한 시간대에 우울이 측정되었다. 우울은 아침에 더 심해지는 등 하루 내에서도 심각도가 변화하는 경향이 있음을 고려할 때(Murray, 2017; Peeters, Berkhof, Delespaul, Rottenberg, & Nicolson, 2006), 하루의 여러 시간대에 걸쳐 수집된 자료는 측정하고자 하는 구인의 특성을 보다 잘 반영한 자료라고 할 수 있다. 다음으로, 센서데이터의 자료 손실을 최소화하였다. 본 연구를 포함한 선행연구들에서 센서데이터는 항시 수집되는 반면, 자기보고 자료는 매우 제한적으로 측정되어(예, 1일 1회), 둘을 비교하기 위해 센서데이터의 하루 평균치를 구하는 등의 데이터 처리를 실시하였다. 평균은 자료의 대푯값으로 기능할 수 있지만, 한편으로는 개별의 자료를 반영하기 어려워 자료의 손실이 있다는 한계점을 지닌다. 이 때, 평균내는 간격을 최소화함으로써 자료의 손실을 줄일 수 있는데, 본 연구에서는 EMA 간격에 맞추어 4시간 당 1개의 값으로 특성정보를 추출함으로써 기존 연구들(최소 24시간)과 비교해 그 간격을 현저히 축소하였다. 본 연구에서는 참가자들의 피로도를 고려하여 1일 4회 설문을 실시하였는데, 추후 이 간격을 더욱 최소화한다면 항시 수집되는 센서데이터의 준거자료로서 더 타당하게 기능할 것으로 예상된다.

본 연구는 센서데이터를 이용한 우울 예측을 위한 초기 시도이다. 비록 작은 수의 대상자와 제

한된 분석방법 등의 한계가 있으나, 우울 증상과 수준에 대한 고려가 있었다는 점에서 관련 연구 분야에 기여하는 바가 있다. 현재까지 센서데이터를 이용한 연구들의 예측률은 80% 정도에 그치는데, 패시브 센싱이 임상 및 연구 장면에서 보다 효용성을 지니기 위해서는 예측률을 높일 필요가 있다. 본 연구에서 시도한 세부 우울 증상 관련 센서데이터의 규명과 우울 수준에 따른 센서데이터의 차이에 대한 고려는 예측률을 높이는 방안 중 일부로 기능할 수 있을 것으로 판단된다. 특히, 본 연구에서 사용된 지표들 중 실제 임상 현장에서 활용할 수 있는 지표들은 다음과 같다. 우선, 증상의 측면에서 각 센서데이터가 서로 다른 증상과 관련 있기에 어느 증상을 측정하는지에 따라 유용한 지표가 달라질 것으로 보인다. 다만, 우울증에서 우울한 기분, 흥미 저하의 두 증상이 특히 중요하다는 점을 고려할 때, 두 증상 모두와 관련이 높은 다음의 증상들이 가장 유용하다고 판단된다: 주요 동작 수, 이동 총 거리, 장소 간 거리의 분산, 방문 장소의 개수, 두 장소간 최장 거리, 집에서의 최장 거리, 평균 심박수. 아울러 우울 수준의 측면에서 또한 모든 지표에서 집단 간 차이가 유의하긴 하였으나, 특히 세 집단간 차이가 모두 유의한 다음의 지표가 수준을 구분하는데 유용할 것으로 판단된다: 휴대폰 기록, 걸음 수, 장소간 거리의 분산, 집에서의 최장거리.

본 연구의 한계점 및 후속 연구에 대한 제언은 다음과 같다. 첫째, 본 연구는 비교적 장기간 스마트기기를 사용해야 한다는 점에서, 기기에 익숙한 대학생 및 대학원생을 대상으로 하였다. 그러나, 대상자의 인구통계학적 변인에 따라 센서데이터와 우울 간 관계에서 차이를 보일 수 있기에, 본

연구 결과의 일반화를 위해서는 다른 인구통계학적 특성을 지닌 집단을 대상으로 하는 추가적인 연구가 필요하다. 둘째, 본 연구에서는 연구가 진행될 당시 양호한 정확도로 구현 가능한 센서 위주로 포함하였는데, 최근에는 기술의 발전에 따라 보다 다양한 센서데이터의 수집이 가능해졌다. 구체적으로, 생리적 지표(예, 피부전도도), 사용자의 기기에 대한 정보(예, WiFi 및 블루투스로 연결된 기기 개수, 배터리 소모량)와 외부 환경에 대한 정보(예, 습도, 온도, 조도)의 종류가 다양해졌으며, 사용자의 기분을 보다 직접적으로 반영하는 음성과 얼굴 표정의 수집이 시도되고 있다(Rohani et al., 2018). 특히, 본 연구에서 유의한 측정에 실패한 두 증상과 관련하여, 사용자의 식욕 및 인지적 측면과 보다 밀접한 관련이 있는 센서를 포함할 필요가 있다. 셋째, 본 연구에서는 데이터의 양과 배터리 소모에 미치는 영향을 고려하여 각 EMA 실시 직전 30분 동안의 센서데이터만이 전송되었다. 총 30일간 1일 4회, 즉 최대 120회 데이터를 측정함으로써 상기 방법으로 인한 한계점을 보완하고자 하였으나 일정 기간의 데이터 손실이 불가피하였다. 추후 연구에서는 기술의 발전을 도모하여 데이터 손실을 최소화하는 것이 필요할 것이다. 넷째, EMA 설문은 6~22시경에 대해서만 실시되어 이외의 시간대에 대한 참가자들의 우울 수준 및 센서데이터 자료를 얻을 수 없었다. 이는 환자의 기상 및 취침 시간 등 개인차 변인에 대해 충분히 고려하지 못한 것으로, 수집된 자료에 편향이 발생할 위험이 있다. 본 연구에서는 기술상의 한계로 인해 연구 결과에 영향을 줄 수 있는 모든 개인차 변인을 포함하지 못했으나 추후 연구에서는 보다 정확한 연구 결과를 위해 이를

고려하고 연구 설계에 반영할 필요가 있다.

마지막으로, 우울의 심각도와 발현 양상은 시간에 따라 변하는 경향이 있으므로(Richards, 2011), 보다 장기간의 연구를 통해 시간에 따른 센서데이터와 우울 간의 관계를 살펴보는 것이 권고된다. 특히, 우울 삽화의 평균 지속 기간이 3~6개월이라는 점을 고려할 때(Furukawa, Kiturama, & Takahashi, 2000; Spijker et al., 2002), 최소 3개월 이상의 연구기간이 타당한 결과를 얻기 위해 필요할 것으로 보인다.

참 고 문 헌

- 성형모, 김정범, 박영남, 배대석, 이선희, 안현의 (2008). 한국어판 백 우울 설문지 2 판의 신뢰도 및 타당도 연구. *생물치료정신의학*, 14(2), 201-212.
- 안제용, 서은란, 임경희, 신재현, 김정범 (2013). 한국어판 우울증 선별도구 (Patient Health Questionnaire-9, PHQ-9)의 표준화 연구. *생물치료정신의학*, 19(1), 47-56.
- 임선영, 이은정, 정성원, 김희철, 정철호, 진태연, 김정범. (2011). 한국판 백 우울 척도 2 판의 타당화 연구. *대한불안학회지*, 7(1), 48-53.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. American Psychiatric Pub.
- Balbin, J. R., Pinugu, J. N. J., Basco, A. J. S., Cabanada, M. B., Gonzales, P. M. V., Marasigan, J. C. C., & Sejera, M. M. (2017). Development of scientific system for assessment of post-traumatic stress disorder patients using physiological sensors and feature extraction for emotional state analysis. *2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, 1-6.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck depression inventory manual (2nd ed.)*. San Antonio, TX: Psychological Corporation.
- Beiwinkel, T., Kindermann, S., Maier, A., Kerl, C., Moock, J., Barbian, G., & Rössler, W. (2016). Using smartphones to monitor bipolar disorder symptoms: a pilot study. *JMIR Mental Health*, 3(1), e2. doi: 10.2196/mental.4560.
- Biel, J. I., Martin, N., Labbe, D., & Gatica-Perez, D. (2018). Bites 'n'Bits: Inferring Eating Behavior from Contextual Mobile Data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(4), 1-33.
- Canzian, L., & Musolesi, M. (2015). Trajectories of depression: Unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. *UbiComp 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 1293 - 1304.
- Coyne, S. M., Stockdale, L., & Summers, K. (2019). Problematic cell phone use, depression, anxiety, and self-regulation: Evidence from a three year longitudinal study from adolescence to emerging adulthood. *Computers in Human Behavior*, 96, 78-84.
- Dinas, P. C., Koutedakis, Y., & Flouris, A. D. (2011). Effects of exercise and physical activity on depression. *Irish Journal of Medical Science*, 180(2), 319-325.
- Farhan, A. A., Lu, J., Bi, J., Russell, A., Wang, B., & Bamis, A. (2016). Multi-view Bi-clustering to Identify Smartphone Sensing Features Indicative of Depression. *Proceedings - 2016 IEEE 1st International Conference on Connected Health: Applications, Systems and Engineering*

- Technologies, CHASE 2016*, 264 - 273.
- Farhan, A. A., Yue, C., Morillo, R., Ware, S., Lu, J., Bi, J., Kamath, J., Russell, A., Bamis, A., & Wang, B. (2016). Behavior vs. introspection: Refining prediction of clinical depression via smartphone sensing data. *2016 IEEE Wireless Health, WH 2016*, 30 - 37.
- Faurholt-Jepsen, M., Busk, J., Dórarinsdóttir, H., Frost, M., Bardram, J. E., Vinberg, M., & Kessing, L. V. (2019). Objective smartphone data as a potential diagnostic marker of bipolar disorder. *Australian & New Zealand Journal of Psychiatry, 53*(2), 119-128.
- First, M. B., Williams, J. B., Karg, R. S., & Spitzer, R. L. (2016). *SCID-5-CV: Structured Clinical Interview for DSM-5 Disorders, Clinician Version*. Washington, DC: American Psychiatric Association Publishing.
- Fried, E. I., & Nesse, R. M. (2015). Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. *BMC Medicine, 13*(1), 1-11.
- Furukawa, T. A., Kiturama, T. & Takahashi, K. (2000). Time to recovery of an inception cohort with hitherto untreated unipolarmajor depressive episodes. *British Journal of Psychiatry, 177*, 331-335.
- Garg, H., Bush, S., & Gappmaier, E. (2016). Associations between fatigue and disability, functional mobility, depression, and quality of life in people with multiple sclerosis. *International Journal of MS Care, 18*(2), 71-77.
- Ghandeharioun, A., Fedor, S., Sangermano, L., Ionescu, D., Alpert, J., Dale, C., Sontag, D., & Picard, R. (2017). Objective assessment of depressive symptoms with machine learning and wearable sensors data. *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, 325-332.
- Hamilton, M. A. X. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology, 6*(4), 278-296.
- Holländare, F., Askerlund, A. M., Nieminen, A., & Engström, I. (2008). Can the BDI-II and MADRS-S be transferred to online use without affecting their psychometric properties?. *Electronic Journal of Applied Psychology, 4*(2), 63-65.
- Huckvale, K., Venkatesh, S., & Christensen, H. (2019). Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. *npj Digital Medicine, 2*(1), 1-11.
- Insel, T. R. (2017). Digital phenotyping: technology for a new science of behavior. *JAMA, 318*(13), 1215-1216.
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry, 167*(7), 748-751.
- Kahveci, A. Y., Alemdar, H., & Ersoy, C. (2015). Sleep quality monitoring with ambient and mobile sensing. *2015 23rd Signal Processing and Communications Applications Conference (SIU)*, 507-510.
- Kemp, A. H., Quintana, D. S., Gray, M. A., Felmingham, K. L., Brown, K., & Gatt, J. M. (2010). Impact of depression and antidepressant treatment on heart rate variability: a review and meta-analysis. *Biological Psychiatry, 67*(11), 1067-1074.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ 9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 87-93.

- Medicine*, 10(9), 606–613.
- Liang, Y., Zheng, X., & Zeng, D. D. (2019). A survey on big data-driven digital phenotyping of mental health. *Information Fusion*, 52, 290–307.
- McDonald, A. D., Sasangohar, F., Jatav, A., & Rao, A. H. (2019). Continuous monitoring and detection of post-traumatic stress disorder (PTSD) triggers among veterans: a supervised machine learning approach. *IISSE Transactions on Healthcare Systems Engineering*, 9(3), 201–211.
- Meegahapola, L., Ruiz-Correa, S., & Gatica-Perez, D. (2020, November). Alone or With Others? Understanding Eating Episodes of College Students with Mobile Sensing. In *19th International Conference on Mobile and Ubiquitous Multimedia* (pp. 162–166). doi.org/10.1016/j.inffus.2019.04.001
- Mehrotra, A., Hendley, R., & Musolesi, M. (2016). Towards multi-modal anticipatory monitoring of depressive states through the analysis of human-smartphone interaction. *UbiComp 2016 Adjunct - Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 1132–1138.
- Mohr, D. C., Zhang, M., & Schueller, S. M. (2017). Personal sensing: understanding mental health using ubiquitous sensors and machine learning. *Annual Review of Clinical Psychology*, 13, 23–47.
- Murray, G. (2007). Diurnal mood variation in depression: a signal of disturbed circadian function?. *Journal of Affective Disorders*, 102(1–3), 47–53.
- Narziev, N., Goh, H., Toshnazarov, K., Lee, S. A., Chung, K. M., & Noh, Y. (2020). STDD: Short-Term Depression Detection with Passive Sensing. *Sensors*, 20(5), 1396. doi:10.3390/s20051396
- Onnela, J. P., & Rauch, S. L. (2016). Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. *Neuropsychopharmacology*, 41(7), 1691–1696.
- Otte, C., Gold, S. M., Penninx, B. W., Pariante, C. M., Etkin, A., Fava, M., Mohr, D. C., & Schatzberg, A. F. (2016). Major depressive disorder. *Nature Reviews Disease Primers*, 2(1), 1–20.
- Park, S. J., Choi, H. R., Choi, J. H., Kim, K. W., & Hong, J. P. (2010). Reliability and validity of the Korean version of the Patient Health Questionnaire-9 (PHQ-9). *Anxiety and Mood*, 6(2), 119–124.
- Peeters, F., Berkhof, J., Delespaul, P., Rottenberg, J., & Nicolson, N. A. (2006). Diurnal mood variation in major depressive disorder. *Emotion*, 6(3), 383–391.
- Peluso, M. A. M., & Andrade, L. H. S. G. D. (2005). Physical activity and mental health: the association between exercise and mood. *Clinics*, 60(1), 61–70.
- Penninx, B. W., Milaneschi, Y., Lamers, F., & Vogelzangs, N. (2013). Understanding the somatic consequences of depression: biological mechanisms and the role of depression symptom profile. *BMC Medicine*, 11(1), 129. doi.org/10.1186/1741-7015-11-129
- Pickett, K., Yardley, L., & Kendrick, T. (2012). Physical activity and depression: A multiple mediation analysis. *Mental Health and Physical Activity*, 5(2), 125–134.
- Place, S., Blanch-Hartigan, D., Rubin, C., Gorrostieta, C., Mead, C., Kane, J., Marx, B., Feast, J., Deckersbach, T., Pentland, A., Nierenberg, A., & Azarbajani, A. (2017). Behavioral indicators on a mobile sensing platform predict clinically validated psychiatric symptoms of mood and anxiety disorders. *Journal of Medical Internet Research*, 19(3), e75. doi: 10.2196/jmir.6678

- Pratap, A., Atkins, D. C., Renn, B. N., Tanana, M. J., Mooney, S. D., Anguera, J. A., & Areán, P. A. (2019). The accuracy of passive phone sensors in predicting daily mood. *Depression and Anxiety, 38*(1), 72-81.
- Raballo, A. (2018). Digital phenotyping: an overarching framework to capture our extended mental states. *The lancet. Psychiatry, 5*(3), 194-195.
- Reinertsen, E., & Clifford, G. D. (2018). A review of physiological and behavioral monitoring with digital sensors for neuropsychiatric illnesses. *Physiological Measurement, 39*(5), 05TR01.
- Richards, D. (2011). Prevalence and clinical course of depression: a review. *Clinical Psychology Review, 31*(7), 1117-1125.
- Rohani, D. A., Faurholt-Jepsen, M., Kessing, L. V., & Bardram, J. E. (2018). Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: systematic review. *JMIR mHealth and uHealth, 6*(8), e165. doi: 10.2196/mhealth.9691
- Saeb, S., Lattie, E. G., Kording, K. P., & Mohr, D. C. (2017). Mobile Phone Detection of Semantic Location and Its Relationship to Depression and Anxiety. *JMIR MHealth and UHealth, 5*(8), e112. doi: 10.2196/mhealth.7297.
- Saeb, S., Lattie, E. G., Schueller, S. M., Kording, K. P., & Mohr, D. C. (2016). The relationship between mobile phone location sensor data and depressive symptom severity. *PeerJ, 2016*(9), 1-15.
- Saeb, S., Zhang, M., Karr, C. J., Schueller, S. M., Corden, M. E., Kording, K. P., & Mohr, D. C. (2015). Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study. *Journal of Medical Internet Research, 17*(7), 1-11.
- Sano, A., Rahman, T., Zhang, M., Ganesan, D., & Choudhury, T. (2020). Mobile Sensing of Alertness, Sleep and Circadian Rhythm: Hardware & Software Platforms. *GetMobile: Mobile Computing and Communications, 23*(3), 16-22.
- Shin, S., Yeom, C. W., Shin, C., Shin, J. H., Jeong, J. H., Shin, J. U., & Lee, Y. R. (2016). Activity monitoring using a mHealth device and correlations with psychopathology in patients with chronic schizophrenia. *Psychiatry Research, 246*, 712-718.
- Spijker, J. A. N., De Graaf, R., Bijl, R. V., Beekman, A. T., Ormel, J., & Nolen, W. A. (2002). Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *The British Journal of Psychiatry, 181*(3), 208-213.
- Tamura, H., Nishida, T., Tsuji, A., & Sakakibara, H. (2017). Association between excessive use of mobile phone and insomnia and depression among Japanese adolescents. *International Journal of Environmental Research and Public Health, 14*(7), 701. doi.org/10.3390/ijerph14070701
- Torous, J., Kiang, M. V., Lorme, J., & Onnela, J. P. (2016). New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. *JMIR mental health, 3*(2), e16. doi: 10.2196/mental.5165
- Torous, J., Onnela, J. P., & Keshavan, M. (2017). New dimensions and new tools to realize the potential of RDoC: digital phenotyping via smartphones and connected devices. *Translational Psychiatry, 7*(3), e1053. doi.org/10.1038/tp.2017.25
- Tron, T., Resheff, Y. S., Bazhmin, M., Peled, A., & Weinshall, D. (2017, November). Real-time schizophrenia monitoring using wearable motion

- sensitive devices. *In International Conference on Wireless Mobile Communication and Healthcare* (pp. 242-249). Springer, Cham.
- Vallée, J., Cadot, E., Roustit, C., Parizot, I., & Chauvin, P. (2011). The role of daily mobility in mental health inequalities: the interactive influence of activity space and neighbourhood of residence on depression. *Social Science & Medicine*, 73(8), 1133-1144.
- Wahle, F., Kowatsch, T., Fleisch, E., Rufer, M., & Weidt, S. (2016). Mobile Sensing and Support for People With Depression: A Pilot Trial in the Wild. *JMIR MHealth and UHealth*, 4(3), e111. doi: 10.2196/mhealth.5960.
- Wang, R., Wang, W., DaSilva, A., Huckins, J. F., Kelley, W. M., Heatherton, T. F., & Campbell, A. T. (2018). Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1), 1 - 26.
- Williamson, J. R., Godoy, E., Cha, M., Schwarzentruher, A., Khorrami, P., Gwon, Y., Kung, H. T., Dagli, C., & Quatieri, T. F. (2016, October). Detecting depression using vocal, facial and semantic communication cues. *In Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge* (pp. 11-18).
- Williamson, G. M., & Shaffer, D. R. (2002). The activity restriction model of depressed affect. *In Physical illness and depression in older adults* (pp. 173-200). MA, Boston: Springer.
- Yue, C., Ware, S., Morillo, R., Lu, J., Shang, C., Bi, J., Kamath, J., Russell, A., Barnis, A., & Wang, B. (2018). Fusing location data for depression prediction. *IEEE Transactions on Big Data*.
- Zimmerman, M., Balling, C., Chelminski, I., & Dalrymple, K. (2018). Understanding the severity of depression: Which symptoms of depression are the best indicators of depression severity?. *Comprehensive Psychiatry*, 87, 84-88.

원고접수일: 2021년 3월 18일

논문심사일: 2021년 3월 26일

게재결정일: 2021년 5월 25일

Investigation of a Passive Sensing Pattern by the Depression Symptom and Level

Seungmin Jung

Department of Psychiatry
Korea University Anam Hospital

Kyong-mee Chung

Department of Psychology
Yonsei University

Passive sensing, which collects behavioral, physiological, social, and environmental indices with smart devices in an objective and automatic manner, is recently being explored as a tool for evaluating mental disorders. Among them, studies on depressive disorder are the most commonly performed. Although the expression patterns of depressive disorder may differ according to the symptoms and levels, studies that take this point into account while examining the relationship with sensor data are very limited. The purpose of this study was to identify sensor data that is highly related to depression symptoms, and to determine whether there is a difference in sensor data according to the level of depression. A total of 64 college students and graduate students were classified into three groups (normal, mildly depressed, and severely depressed) according to the level of depression. For a total of 30 days, self-reported data on 9 symptoms of depressive disorder were repeatedly collected 4 times a day. At the same time, a total of 8 sensor data were collected throughout the day using a smartphone and a smartwatch, and a total of 14 features were extracted as 4 values per day. According to the results of this study, different features were found to have a significant relationship with each symptom. Some symptoms were found to be related with all features, some symptoms were related with only partial features, and some symptoms were not related with any features. In addition, it was found that there were significant differences in all features according to the level of depression, and detailed patterns were different for each feature. Implications and limitations of the current study and directions for future research were further discussed.

Keywords: passive sensing, depression, symptom, level