KMAN PUBLISHING INSTITUTE
Journal Website

Article history:
Received 20 January 2026
Revised 15 March 2026
Accepted 19 March 2026
Published online 01 April 2026

Journal of Personality and Psychosomatic Research

Open Peer-Review Report

E-ISSN: 3041-8542

Personality-Based Digital Phenotyping of Psychosomatic Health Using Machine Learning

Lena. Hoffmann[1*]

[1] Affiliation: Department of Cognitive Psychology, Heidelberg University, Heidelberg, Germany

*** Corresponding author email address**: lena.hoffmann@psychologie.uni-heidelberg.de

Editor	Reviewers
Ali Aghaziarati Department of Psychology and Counseling, KMAN Research Institute, Richmond Hill, Ontario, Canada aliaghaziarati@kmanresce.ca	**Reviewer 1:** Ikeagwu Ugbeh-Johnson School of Public Health and Community Medicine,University of New South Wales, Sydney, Australia. Email: upbeh-johnsonikeagwy@gmail.com **Reviewer 2:** Daymar Avilés-Santa National Institute of Psychiatry Ramón de la Fuente Muñiz, Sub-directorate of Clinical Research, Mexico City, Mexico. daymaraviles-santa10@gmail.com

1. Round 1

1.1. Reviewer 1

Reviewer:

The sampling strategy and inclusion criteria are described clearly, but there is insufficient information on the representativeness of the sample and possible selection bias: given recruitment via university mailing lists and online platforms, it is likely that highly educated and digitally engaged individuals are overrepresented; please provide more detailed descriptive statistics on education, employment, and regional distribution, discuss how this may limit generalizability to populations with different socioeconomic and digital literacy profiles, and consider whether weighting or sensitivity analyses are warranted.

The interpretation of feature importance, particularly the emphasis on neuroticism and nighttime smartphone activity as key predictors, risks being overstated unless you clarify the method used (e.g., Gini importance, permutation importance, SHAP values) and its known biases; I recommend reporting uncertainty around importance estimates, illustrating the directionality of effects (e.g., with partial dependence plots or SHAP summary plots), and explicitly acknowledging that "importance" does not directly imply causality or even linear monotonic relationships.

Although ethical and data protection aspects are mentioned, given the sensitivity of continuous smartphone monitoring and psychosomatic health prediction, the manuscript would benefit from a more extensive ethics section: please elaborate on consent procedures (especially regarding passive sensing), options for opting out of specific data streams, participants'

awareness of the inferential power of the models (e.g., predicting health risk), data retention policies, and the potential risks of misuse (e.g., by employers or insurers), along with safeguards for responsible deployment.

The reporting of results for psychometric measures and digital behaviors is generally clear, but the manuscript would be strengthened by more rigorous statistical reporting and structure: for each main correlation and group difference, please report exact p-values, confidence intervals, and effect sizes; provide full table captions that describe what is adjusted/unadjusted; clarify whether multiple comparison corrections were applied; and consider presenting a concise multivariate model (e.g., regression baseline) alongside the ML models to anchor the findings in more traditional statistical terms.

Authors revised the manuscript and uploaded the document.

1.2. Reviewer 2

Reviewer:

While the description of digital phenotyping variables is commendably rich, the operationalization of key constructs (e.g., "sleep-related phone inactivity periods" as proxy sleep duration, "mobility-related indicators" as proxy physical activity) should be described with more technical precision: please specify exact computation rules (e.g., minimum inactivity window to count as "night," handling of shift workers, thresholds and time windows for mobility variability), and discuss the known limitations of these proxies relative to gold-standard measures such as actigraphy or GPS-based mobility indices.

The machine learning pipeline is presented in broad strokes, but for a scientific audience the analytic transparency is currently insufficient: important modeling details are missing, including the handling of class imbalance in psychosomatic risk categories, the hyperparameter tuning strategy (e.g., grid search vs. randomized search and parameter ranges), the exact definition of the predicted "low versus elevated" psychosomatic categories (PHQ-15 cutoffs and whether DASS-21 was also incorporated into the target), and whether nested cross-validation was used to avoid optimistic bias in model selection.

The performance metrics of the Gradient Boosting model (accuracy = 0.82, AUC = 0.87) are promising, but the clinical and practical significance of these results remains underdeveloped: you should report class-specific sensitivity and specificity, positive/negative predictive values, and confusion matrices for the chosen risk thresholds, and then explicitly discuss whether these levels of performance would be adequate for screening, risk stratification, or intervention triage in realistic healthcare or workplace settings.

Authors revised the manuscript and uploaded the document.

2. Revised

Editor's decision: Accepted.
Editor in Chief's decision: Accepted.

JPPR
Journal of Personality and Psychosomatic Research
E-ISSN: 3041-8542