Exploring Wearable-Based Detection of Compulsive Handwashing in a Non-Controlled Setting: A Case Study

Kristina Kirsten
Digital Health - Connected Healthcare, Hasso Plattner Institute, University of Potsdam
Potsdam, Germany

Robin Burchard
Ubiquitous Computing, University of Siegen
Siegen, Germany

Amatya Mackintosh
Faculty of Psychology, University of Basel
Basel, Switzerland

Marcel Miché
Faculty of Psychology, University of Basel
Basel, Switzerland

Dorothée Bentz
Faculty of Psychology, University of Basel
Basel, Switzerland

Klaus Bader
Center for Psychosomatics and Psychotherapy, University Psychiatric Clinics (UPK)
Basel, Switzerland

Jannis Behr
Center for Psychosomatics and Psychotherapy, University Psychiatric Clinics (UPK)
Basel, Switzerland

Roselind Lieb
Faculty of Psychology, University of Basel
Basel, Switzerland

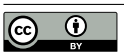
Kristof Van Laerhoven
Ubiquitous Computing, University of Siegen
Siegen, Germany

Bert Arnrich
Digital Health - Connected Healthcare, Hasso Plattner Institute, University of Potsdam
Potsdam, Germany

Karina Wahl
Faculty of Psychology, University of Basel
Basel, Switzerland

Abstract

Obsessive-compulsive disorder (OCD) is characterized by repetitive behaviors such as compulsive handwashing, yet there is limited sensor-based research on detecting and characterizing such behaviors in real-world settings. This case study presents an unobtrusive wearable system that combines smartwatch-based motion data and location data to monitor handwashing activities in a psychiatric ward. Over 63 hours of data were collected from one participant, capturing 41 labeled handwashing acts. We explore behavioral patterns, correlate motion features with self-reported urge and distress levels, and demonstrate the feasibility of detecting handwashing using machine learning. A Random Forest model achieved strong performance (F1 = 0.710), while an LSTM model demonstrated high specificity (0.995) and moderate sensitivity (0.606). Additionally, we performed binary classification between compulsive and routine handwashing, achieving an F1 score of 0.94. These findings support future real-world research on wearable-based OCD monitoring and promote the integration of wearables into digital phenotyping for mental health.



ISWC '25, Espoo, Finland

ACM ISBN 979-8-4007-1481-8/2025/10
https://doi.org/10.1145/3715071.3750434

CCS Concepts

• **Human-centered computing** → **Personal digital assistants**; • **Applied computing** → Consumer health; *Health care information systems*; **Psychology**; • **General and reference** → *Design.*

Keywords

Digital Mental Health, Wearables, HAR, OCD

ACM Reference Format:
Kristina Kirsten, Robin Burchard, Amatya Mackintosh, Marcel Miché, Dorothée Bentz, Klaus Bader, Jannis Behr, Roselind Lieb, Kristof Van Laerhoven, Bert Arnrich, and Karina Wahl. 2025. Exploring Wearable-Based Detection of Compulsive Handwashing in a Non-Controlled Setting: A Case Study. In *Proceedings of the 2025 ACM International Symposium on Wearable Computers (ISWC '25), October 12–16, 2025, Espoo, Finland.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3715071.3750434

1 Introduction

Smart devices are increasingly used in daily life, including preventive healthcare and rehabilitation [1, 6]. Real-time data collection offers a holistic view of an individual's condition by capturing routines, context, and symptom changes. Wearables can enhance diagnosis, personalize treatment, and support relapse prevention in mental health. However, most studies focus on depression and anxiety [4], likely due to their high prevalence and larger cohorts. Still, less-studied disorders may benefit equally from such monitoring. One such disorder is obsessive-compulsive disorder (OCD), marked

by intrusive thoughts (obsessions) and repetitive behaviors (compulsions). Despite its relevance, it is underrepresented in wearable research, and no public dataset contains motion data from OCD patients. Our study addresses this gap by focusing on sensor-based detection and characterization of compulsive behaviors. We present findings from a case study with one participant, as recruitment in this context is especially challenging. Still, analyzing a single patient allows us to test hypotheses and refine methods. Conducted in a psychiatric ward, the study collects sensor data on (compulsive) handwashing and daily activity prior to treatment. With more participants, we aim to gain broader insights and contribute a unique dataset to the fields of activity recognition and psychology.

2 Background

OCD includes a wide range of compulsive behaviors, many of which align with tasks studied in human activity recognition (HAR) due to their repetitive, structured nature, making them suitable for sensor-based detection. However, symptom variability complicates the creation of comprehensive datasets and generalizable models. To address this, we focus initially on (compulsive) handwashing, one of the most common OCD behaviors, as a starting point to explore its characteristics and compare them to existing research. Nonetheless, digital mental health research in real-world, semi-supervised settings poses additional challenges, including recruitment, data quality, and ethical considerations in clinical populations.

2.1 Related Work

Handwashing detection as part of general hand hygiene has been widely studied [10], using wearable motion sensors, computer vision, and machine learning. However, most work was conducted in controlled settings with scripted procedures such as WHO guidelines [5, 8, 11], lacking evaluation in unobtrusive, real-world deployments. Digital mental health research on OCD is emerging, but integrating consumer-grade wearables remains limited, particularly in real-world, unsupervised settings. A scoping review [3] summarized wearable and smartphone-based tools for assessing, monitoring, or treating OCD, highlighting the evolving nature of this research. A pilot study by Olesen et al. [7] showed the feasibility of predicting OCD events in children and adolescents using wearables, supporting their potential for in-the-wild monitoring.

Our previous work demonstrated that enacted compulsive handwashing behavior can be distinguished in controlled settings using wearable data and machine learning [2, 9]. The present study builds on this by transferring our system to a non-controlled clinical setting to gain insights into real-world compulsive behaviors.

2.2 Study Description

In our ongoing study, approved by the ethics committee of Northwestern Switzerland (ID: 2024-00601), we observe compulsive handwashing in patients undergoing treatment for OCD in a psychiatric ward. Data is collected during the week before exposure and response prevention (ERP) therapy to establish a ground truth dataset for later analysis. Participants are enrolled shortly after admission and instructed to wear a provided smartwatch daily for one week, from morning to evening. They log each handwashing episode on the watch and answer two brief questions about their emotional state (urge and distress). To preserve ecological validity, interaction with the device is minimized. Because using the watch before handwashing could influence behavior, participants are not asked to mark the start of the wash. A supervised trial handwashing session is conducted beforehand to ensure participants understand the procedure, verify system functionality, and provide a standardized reference for analysis.

2.3 Technological Study Setup

Each participant is equipped with a consumer-grade smartwatch[1] that operates independently via an integrated SIM card. The watch runs a custom app used to record data, mark handwashes as compulsive or routine, and answer two questions (urge and distress levels on a 5-point Likert scale) after each event. Participants also rate their overall symptom severity each evening. A compulsive handwash is defined as driven by a strong urge, often due to contamination fears or the need to prevent harm, while routine handwashing occurs without such urges, for example, before meals. It remains to be explored how clearly patients distinguish between these types. Bluetooth beacons placed at sinks provide location context by broadcasting signals used to estimate proximity. This setup avoids audio or video sensors, preserving privacy, especially important in psychiatric settings.

3 Sensor-Based Behavioral Analysis

The dataset collected in this study offers a unique opportunity to examine compulsive behavior through wearable sensing in a real-world (clinical) environment. Even beyond its relevance to OCD, the resulting dataset provides valuable contributions to the broader field of handwashing detection, as it was collected in a naturalistic, non-controlled setting without reliance on video or audio data.

Given the clinical sensitivity of the disorder and the complexity of compulsive actions, the data must be interpreted with care. The continuous, minimally invasive setup enables unobtrusive behavior monitoring while preserving the authenticity of participants' actions. This opens the possibility for novel insights into the relationship between objective motion data and subjective questionnaire data, explored through analyses of event patterns, correlations with self-reported measures, and machine learning-based classification.

3.1 Multimodal Data Exploration

Early data exploration is crucial to identify setup issues, assess data quality, and guide hypothesis testing, especially in psychiatric settings where OCD symptoms and limited participant availability may affect recruitment and device use. For our first participant, 63.88 hours of data were collected, including 41 marked handwashing acts (33 compulsive, eight routine). On average, compulsive episodes lasted 2.21 minutes; routine episodes, 1.71 minutes. The participant described the week as good and noted that distinguishing between compulsive and routine handwashing was not always straightforward. We also observed partial beacon data loss for several hours, which provided valuable feedback to improve the app before wider deployment. Mean urge and distress ratings showed clear differences in subjective experience: compulsive handwashing (urge: 2.48, distress: 2.18) was linked to greater emotional intensity than

[1] Google Pixel Watch 2 LTE

routine acts (urge: 0.62, distress: 0.50). Figure 1 shows representative multimodal sensor data from a handwashing event, combining motion and proximity signals to support activity segmentation. The beacon signal contextualizes motion data by indicating proximity to the sink, enhancing interpretability. In the context of compulsive handwashing, we are particularly interested in episode duration, as the behavior typically involves prolonged activity. Additionally, the data allows exploration of influencing and individual factors such as time of day, day of the week, subjective urge and distress levels, and motion frequency and intensity.

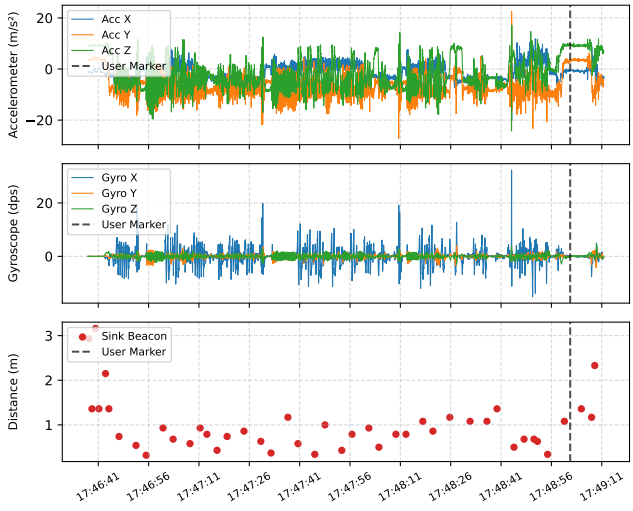


Figure 1: Accelerometer and gyroscope data from a handwash. The dashed line marks the participant's input; the lower subplot shows beacon proximity for activity segmentation.

3.2 Activity Pattern Analysis

To explore the feasibility of training classifiers for handwashing detection, we compared the supervised handwash to one recorded during the data collection phase using two similarity measures. Dynamic time warping (DTW) on raw inertial measurement unit (IMU) signals showed very low similarity, indicating substantial global differences in activity structure. In contrast, a windowed cosine similarity approach with 2-s non-overlapping windows revealed high local similarity (mean: 0.85, max: 0.98). Features were extracted from each window using statistical descriptors (mean, standard deviation, energy, and entropy) across all IMU axes. These results suggest that while full handwashing episodes may differ in duration and global shape, they often contain short, repeated motion patterns suitable for training machine learning models.

3.3 Motion Data and Self-Report Correlations

We examined how self-reports relate to motion data during handwashing, testing whether (1) episode duration and (2) motion features (frequency and energy) correlate with urge and distress levels. We analyzed 41 handwashing acts. For hypothesis (1), we observed moderate positive correlations between the duration and self-reported values for compulsive episodes: $r = 0.60$ for urge and $r = 0.59$ for distress. In contrast, routine episodes showed no meaningful correlation, suggesting that duration is more strongly linked to subjective experience in compulsive contexts. For hypothesis (2), energy-based features, especially from the gyroscope, showed the strongest correlations. *gyro_x_energy* and *gyro_mag_energy* correlated with both urge ($r = 0.59/0.61$) and distress ($r = 0.62/0.62$), while *acc_mag_energy* was also moderately linked (urge: $r = 0.58$; distress: $r = 0.55$). Frequency features had weak or negative associations. The findings suggest that movement intensity, particularly rotational energy, may reflect emotional arousal during compulsions. However, as the analysis is based on a single participant, these results should be interpreted as preliminary and exploratory. Further data from additional participants will be necessary to assess the generalizability and robustness of these findings.

3.4 Machine Learning Classification

We evaluated two approaches for binary classification of handwashing versus other activities (NULL) using 1-s sliding windows with 50% overlap. To address class imbalance, we applied random oversampling during training. Distance and beacon identity were included as features when available, though some beacon data were missing for the participant. Stratified 5-fold cross-validation was used to preserve class distributions. First, we trained a Random Forest (RF) classifier with 100 trees using statistical features per window (mean, std, min, max, and signal energy per axis). It achieved an F1 score of 0.710, sensitivity of 0.653, specificity of 0.996, and area under the precision-recall curve (AUPRC) of 0.750, indicating strong performance even on small datasets. Second, we implemented an Long Short-Term Memory (LSTM) model processing raw multivariate sensor sequences, outputting binary predictions via a fully connected head. Despite limited data, it achieved an F1 score of 0.667, sensitivity of 0.606, specificity of 0.995, and AUPRC of 0.677. While overall performance was lower than the RF, the LSTM showed high precision and low false-positive rates, suggesting it effectively captures temporal dynamics. These findings indicate that feature-based models are more robust in small-sample settings, while larger datasets may better support temporal models. Finally, we classified compulsive vs. routine handwashing using only labeled handwash segments. The RF achieved strong results for compulsive handwashing (F1 = 0.94, sensitivity = 0.98), though specificity was lower (0.40). Despite data rebalancing, the small number of routine instances limits conclusions.

4 Conclusion & Outlook

This case study demonstrates the feasibility of using consumer-grade wearables and Bluetooth beacons to detect compulsive handwashing in a psychiatric setting. Combining motion and location data allowed us to segment events and relate sensor patterns to self-reported urge and distress levels. Our machine learning results suggest that feature-based models perform robustly in small-sample settings, while temporal models show potential for capturing behavioral dynamics. These findings lay the basis for future research on model refinement, validation across participants, and clinical integration in the context of OCD. Future work will explore physiological signals like heart rate and skin conductivity, as we expect them to reflect stress during compulsions and further strengthen our multimodal approach.

References

[1] Justin Albert, Lin Zhou, Kristina Kirsten, Nurcennet Kaynak, Torsten Rackoll, Tim Walz, David Weese, Rok Kos, Alexander Heinrich Nave, and Bert Arnrich. 2025. Using Wearable Sensors in Stroke Rehabilitation. In *Sensor-Based Activity Recognition and Artificial Intelligence: 9th International Workshop, IWOAR 2024, Potsdam, Germany, September 26–27, 2024, Proceedings* (Potsdam, Germany). Springer-Verlag, Berlin, Heidelberg, 277–282. https://doi.org/10.1007/978-3-031-80856-2_19

[2] Robin Burchard, Philipp M Scholl, Roselind Lieb, Kristof Van Laerhoven, and Karina Wahl. 2022. WashSpot: real-time spotting and detection of enacted compulsive hand washing with wearable devices. In *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers*. 483–487.

[3] Adam C Frank, Ruibei Li, Bradley S Peterson, and Shrikanth S Narayanan. 2023. Wearable and mobile technologies for the evaluation and treatment of obsessive-compulsive disorder: scoping review. *JMIR Mental Health* 10 (2023), e45572.

[4] Nuno Gomes, Matilde Pato, Andre Ribeiro Lourenco, and Nuno Datia. 2023. A survey on wearable sensors for mental health monitoring. *Sensors* 23, 3 (2023).

[5] Hong Li, Shishir Chawla, Richard Li, Sumeet Jain, Gregory D Abowd, Thad Starner, Cheng Zhang, and Thomas Plötz. 2018. Wristwash: towards automatic handwashing assessment using a wrist-worn device. In *Proceedings of the 2018 ACM international symposium on wearable computers*. 132–139.

[6] Lin Lu, Jiayao Zhang, Yi Xie, Fei Gao, Song Xu, Xinghuo Wu, Zhewei Ye, et al. 2020. Wearable health devices in health care: narrative systematic review. *JMIR mHealth and uHealth* 8, 11 (2020), e18907.

[7] Kristoffer Vinther Olesen, Nicole Nadine Lønfeldt, Sneha Das, Anne Katrine Pagsberg, and Line Katrine Harder Clemmensen. 2023. Predicting obsessive-compulsive disorder events in children and adolescents in the wild using a wearable biosensor (wrist angel): Protocol for the analysis plan of a nonrandomized pilot study. *JMIR research protocols* 12, 1 (2023), e48571.

[8] Sirat Samyoun, Sudipta Saha Shubha, Md Abu Sayeed Mondol, and John A Stankovic. 2021. iWash: A smartwatch handwashing quality assessment and reminder system with real-time feedback in the context of infectious disease. *Smart Health* 19 (2021), 100171.

[9] Karina Wahl, Philipp Marcel Scholl, Silvan Wirth, Marcel Miché, Jeannine Häni, Pia Schülin, and Roselind Lieb. 2022. On the automatic detection of enacted compulsive hand washing using commercially available wearable devices. *Computers in Biology and Medicine* 143 (2022), 105280.

[10] Chaofan Wang, Weiwei Jiang, Kangning Yang, Difeng Yu, Joshua Newn, Zhanna Sarsenbayeva, Jorge Goncalves, and Vassilis Kostakos. 2021. Electronic Monitoring Systems for Hand Hygiene: Systematic Review of Technology. *J Med Internet Res* 23, 11 (24 Nov 2021), e27880. https://doi.org/10.2196/27880

[11] Yiyuan Zhang, Tianwei Xue, Zhenjie Liu, Wei Chen, and Bart Vanrumste. 2021. Detecting hand washing activity among activities of daily living and classification of WHO hand washing techniques using wearable devices and machine learning algorithms. *Healthcare technology letters* 8, 6 (2021), 148–158.