Utilizing Wearable Technology to Characterize and Predict Post-Exertional Malaise Crashes across Post-COVID Syndrome and Chronic Inflammatory Conditions: Study Protocol of the prospective observational study ‘U-WaTCH’

Viktoria Lampe

viktoria.lampe@med.uni-goettingen.de

University Medical Center Göttingen

Tim Riester

University Medical Center Göttingen

Lena-Marie Theil

University Medical Center Göttingen

Priyanka Dey

University Medical Center Göttingen

Philipp Wieder

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)

Frank Müller

University Medical Center Göttingen

Dominik Schröder

University Medical Center Göttingen

Tim Schmachtenberg

University Medical Center Göttingen

Eva Hummers

University Medical Center Göttingen

Eva Maria Noack

University Medical Center Göttingen

Sigrun May

Ostfalia University of Applied Science

Frank Klawonn

Ostfalia University of Applied Science

Georg Martin Behrens

Hannover Medical School

Rebecca Rubin
Hannover Medical School
Alexandra Dopfer-Jablonka
Hannover Medical School
Sandra Steffens
Hannover Medical School
Christine Happle
Hannover Medical School

Study protocol

Keywords: Post-COVID-Syndrome, inflammatory diseases, Post-Exertional Malaise, fatigue, crash, machine learning

Posted Date: April 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-7138872/v1

Additional Declarations: No competing interests reported.

Utilizing Wearable Technology to Characterize and Predict Post-Exertional Malaise Crashes across Post-COVID Syndrome and Chronic Inflammatory Conditions: Study Protocol of the prospective observational study 'U-WaTCH'

Viktoria Lampe[1*], Tim Riester[1*], Lena-Marie Theil[1], Priyanka Dey[1,2], Philipp Wieder[2], Frank Müller[1,3], Dominik Schröder[1], Tim Schmachtenberg[1], Eva Hummers[1], Eva Maria Noack[1], Sigrun May[4], Frank Klawonn[4,5], Georg Martin Behrens[6,7], Rebecca Rubin[6], Alexandra Dopfer-Jablonka[6,7], Sandra Steffens[8], Christine Happle[9,10]

[1] Department of General Practice, University Medical Center Göttingen, Göttingen, Germany

[2] Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG), Göttingen, Germany

[3] Department of Family Medicine, College of Human Medicine, Michigan State University, Grand Rapids, MI, USA

[4] Department of Computer Science, Ostfalia University of Applied Science, Wolfenbüttel, Germany

[5] Biostatistics Research Group, Helmholtz Centre for Infection Research, Braunschweig, Germany

[6] Department of Rheumatology and Immunology, Hannover Medical School, Hannover, Germany

[7] German Center for Infection Research (DZIF), partner site Hannover-Braunschweig, Germany

[8] Deans' office, Curriculum Development, Hannover Medical School, Hannover, Germany

[9] Department of Pediatric Pulmonology, Allergology, and Neonatology, Hannover Medical School, Hannover, Germany

[10] Center for Individualized Infection Medicine (CiiM), Hannover, Germany

* Shared first authorship.

Corresponding Author: Viktoria Lampe
Department of General Practice, University Medical Center Göttingen, Humboldtallee 38, 37073 Göttingen, Germany
E-mail address: viktoria.lampe@med.uni-goettingen.de

Abstract

Background: Fatigue and post-exertional malaise (PEM) are defining symptoms of Post-COVID Syndrome (PCS) but are also common among rheumatic conditions. PEM is a severe symptom exacerbation following minimal physical or cognitive exertion, often described by affected people as "crashes" with rapid increase of fatigue and reduced physical resilience lasting hours to weeks. Current approaches to PEM management rely on subjective self-monitoring, but objective, real-time PEM prediction does not exist. This study aims to develop and validate machine learning algorithms using continuous wearable sensor data to predict individual PEM crashes, enabling personalized PEM intervention strategies.

Methods: We will conduct a prospective observational study recruiting 300 participants across three groups (n=100 each): 1) PCS patients with PEM, 2) patients with inflammatory rheumatic conditions experiencing PEM, and 3) healthy controls. Participants will use the Apple Watch SE or other compatible wearable devices for continuous monitoring over up to 180 days. The initial data recording period for all groups is 90 days and, whereas for PEM patients, a second period of 90 days is added to allow testing of the detection algorithm. We will collect multi-modal data including heart rate variability, activity patterns, sleep metrics, environmental exposure, and patient-reported outcomes. Advanced machine learning algorithms (Cox regression, survival random forests, neural networks) will be trained to predict crash occurrence. A proof-of-concept "PEM-protection mode" will be integrated into an established no-Code mHealth application.

Discussion: This study represents the first systematic attempt to leverage continuous wearable monitoring and machine learning for PEM prediction in PCS and rheumatic diseases. U-WaTCH is planned to lead to digital interventions, that could transform PEM management from reactive symptom reporting to proactive crash prevention, ultimately improving quality of life for people affected by Long COVID and other conditions.

Trial Registration: DRKS00036301 (German Clinical Trials Register, registered on March 27, 2025).

Keywords: Post-COVID-Syndrome, inflammatory diseases, Post-Exertional Malaise, fatigue, crash, machine learning

Background

The Post-COVID-Syndrome (PCS) is characterized by a variety of symptoms that persist after the acute infection – depending on the definition for several weeks or months after the initial infection [1]. PCS significantly affects the quality of life and presents a challenge to healthcare systems globally [2–5]. According to the World Health Organization (WHO) and other studies, approximately six percent of people continue to experience symptoms four weeks after an acute COVID-19 infection, with 80-90% of these individuals reporting symptoms of post-exertional malaise (PEM) [6–8]. The enduring symptoms characteristic of PCS include persistent fatigue, dyspnea, and cognitive dysfunction [9], which can worsen following physical or cognitive activity, an indicator of PEM [10–12]. PEM occurs in PCS and other diseases with chronic fatigue such as the myeloid encephalitis/ chronic fatigue syndrome (ME/CFS) of unknown origin [13] and other chronic inflammatory conditions, such as rheumatoid arthritis or Sjögren's Syndrome [14]. PEM significantly impacts the quality of life and treatment demand of affected individuals [15].

People with PEM are at risk of over-exertion, even during everyday activities such as shopping or social activities, which can trigger "crashes" characterized by a rapid increase of fatigue and reduced physical resilience [13]. Such crashes can occur within a few hours to a day following a triggering activity and may last from hours to weeks [13]. Crashes in PCS patients are typically triggered by physical activities, cognitive tasks, social interactions, self-care activities, and insufficient sleep [15, 16]. Interestingly, in people with inflammatory rheumatic diseases, research indicates that crashes often originate from different mechanisms, with mental rather than physical exertion being the predominant trigger [17]. Interestingly, in people with inflammatory rheumatic diseases, research indicates that crashes often originate from different mechanisms, with mental rather than physical exertion being the predominant trigger [15].

Although many triggers for PEM are known, they largely differ between individuals and PEM episodes cannot be reliably predicted [18]. Consequently, patients limit themselves in their daily activities and

responsibilities, often relying on the support of others. This situation imposes a significant burden not only on the affected individuals but also on families and society as a whole.

Despite distinct clinical presentations, comparative analyses of PEM in PCS vs. rheumatic conditions have been notably absent from current research. Most existing studies examine these conditions in isolation, without addressing shared symptoms or triggers.

To reduce the enormous burden of PEM for affected individuals, predicting crashes and identifying PEM triggers and exertion thresholds can be a crucial step toward developing personalized management strategies. PEM prediction could empower patients to navigate daily activities while minimizing symptom exacerbation, ultimately improving their activity, social participation, and quality of life.

Thus far, there has been a lack of real-time data collection that could elucidate temporal dynamics of crash development and recovery. This represents a significant research gap limiting our ability to develop tailored interventions for different patient populations experiencing PEM.

The here presented study aims at collecting wearable data on parameters such as heart rate, physical activity, and sleep in people with and without PEM and to identify patterns that can be used to predict PEM episodes. The specific objectives of the U-WaTCH study (Utilizing WearAble TeCHnology for Enhanced Monitoring and Management of Long COVID) are to (a) identify trigger factors for crashes in PEM, (b) develop machine learning algorithms to predict individual or group-specific exertion thresholds, (c) conceive a pilot intervention of a wearable-based “PEM-protection-mode” that can be used to warn participants if they exceed exertion thresholds.

Methods

Study design and Setting

The U-WaTCH study is conducted as an observational study at Hannover Medical School (MHH) and the University Medical Center Göttingen (UMG) and enrolls individuals diagnosed with PCS associated PEM, patients with a chronic inflammatory condition associated PEM, and healthy controls. We plan to start including participants consecutively starting from summer 2025. The study region located in Northern Germany encompasses the greater Hannover area (approximately 1.1 million residents [19] and southern Lower Saxony including Göttingen (approximately 330,000 [19] residents, representing both urban and rural populations.

A key element of this study is a prospective observational cohort study that aims to use, among other things, the established no-code mHealth platform CIAS-EU (cias-app.eu) [20] and Apple watches (Apple Inc., Cupertino, CA) to continuously collect health and activity data and patient-reported outcomes from three observational groups with the sample size ratio 1:1:1 (1) PEM in patients with PCS, (2) PEM in patients with inflammatory rheumatic conditions, and (3) healthy controls. Figure 1 indicates the study's timeline.

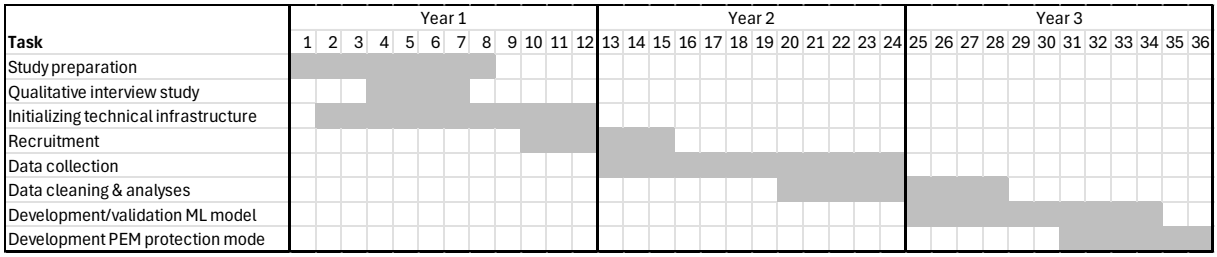

Figure 1: U-WaTCH Study Overview.

Study consortium and roles

The study is conducted by a multidisciplinary consortium: the Department of Rheumatology and Immunology at Hannover Medical School (MHH) serves as coordinator, leading project management and recruitment center for PCS patients and rheumatological patients. The Department of General Practice at University Medical Center Göttingen (UMG) oversees recruitment activities, manages the Citizen's Advisory Board, and contributes expertise in patient-centered digital health interventions. The Gesellschaft für wissenschaftliche Datenverarbeitung (GWDG) provides essential technical infrastructure, maintains the CIAS-EU platform, and ensures data security with their ISO 27001:2013 certified servers. The Ostfalia University of Applied Sciences Department of Computer Science contributes expertise in machine learning and biostatistics, developing and refining the predictive models for PEM crash identification and analysis.

Recruitment

We employ a convenience sampling strategy with consecutive enrolment over a 12-month period. Different recruitment strategies for the respective groups are applied:

Persons in the PCS associated PEM group (1) are recruited through the DEFEAT-Corona observational study platform with approximately 3,000 registered PCS patients [21], specialist PCS outpatient clinics at both university hospitals UMG and MHH, the regional primary care research practice network (Forschungspraxennetzwerk), institutional websites, and social media platforms.

Participants with chronic inflammatory condition associated PEM (2) are recruited through the Cooperative Regional Rheumatology Center Lower Saxony and its network of affiliated practices.

Healthy controls (3) are recruited using a participant referral approach, where enrolled participants from groups 1 and 2 suggest friends, siblings, or work colleagues who might serve as controls. This strategy typically results in lower dropout rates and higher study motivation due to personal

connections. While not formally matched, this approach naturally yields controls with similar demographic characteristics to the patient groups, improving comparability without explicit sampling constraints.

The study targets n=100 participants per group (total n=300). Monthly monitoring of recruitment progress allows for strategic adjustments. If enrolment falls below projected targets, recruitment will be expanded to include additional channels such as patient organizations, self-help groups, and regional media channels and social media. All participants must be able to attend a visit at an inclusion facility for enrolment and device setup.

Potential participants are asked to complete a digital pre-screening survey to assess compliance with initial eligibility criteria. Individuals meeting the pre-screening criteria can schedule a telephone interview with the study personnel. During this call, appointments for enrollment are scheduled. At enrollment, study staff will conduct a final eligibility assessment, obtain written informed consent, and provide comprehensive training on Apple Watch usage and study procedures.

The study provides Apple watch SE (Series 2) devices equipped with eSIM cards and data plans free if charge to participants, ensuring that they incur no additional expenses for taking part in the study. Participants who own an iPhone can pair the study-provided Apple watch with their device and participate by installing the study app, which handles secure data collection and transmission to the study database. For participants without personal iPhone, the study provides a loan device for the duration of their participation. All participants receive 100€ in compensation for their participation.

Eligibility Criteria

Group 1: PCS with PEM

Inclusion criteria are (1) Participants aged 18 years or older, (2) diagnosed PCS (ICD-10 U09.9) or persistent symptoms following a confirmed COVID-19 infection (as determined by PCR or antigen test) lasting more than three months that cannot be explained by other conditions and have first appeared

in association with a COVID-19 infection (following the WHO PCS definition [22]), (3) PEM indicated by the DePaul Symptom Questionnaire-PEM (DSQ-PEM) [23, 24], with thresholds following the National Institute of Neurological Disorders and Stroke Common Data Elements PEM Working Group (at least one of the items 1-5 with ≥2 on both the severity and frequency scales) [25], and (4) Bell Chronic Fatigue Immune Dysfunction Syndrome (CFIDS) disability scale [26] score in the range of 30 to 70.

Exclusion criteria are (1) fatigue due to other conditions (including active cancer or ongoing cancer treatment, chronic inflammatory or autoimmune diseases such as rheumatoid arthritis, systemic lupus erythematosus, inflammatory bowel disease, or multiple sclerosis, severe or symptomatic anemia requiring treatment, chronic infections including Human immunodeficiency virus (HIV), hepatitis B/C, or tuberculosis, primary adrenal insufficiency including Addison's disease, severe organ dysfunction such as heart failure NYHA class III-IV, chronic kidney disease stage 4-5, or end stage liver cirrhosis, neurological conditions associated with significant fatigue including Parkinson's disease or myasthenia gravis), (2) other poorly controlled chronic conditions where fatigue is a predominant symptom (including uncontrolled depression with the Patient Health Questionnaire-9 (PHQ-9) scores ≥15, untreated or poorly controlled hypothyroidism with TSH >10 mIU/L, diabetes with HbA1c >10%, or fatigue attributed to poor glycemic control, untreated obstructive sleep apnea syndrome (OSAS)], and severe nutritional deficiencies, drug abuse, use of medications causing significant fatigue as a primary side effect), (3) lack of capacity to consent, (4) limited German language proficiency, and (5) insufficient technical skills regarding wearable or mobile phone handling.

Group 2: Chronic Inflammatory Diseases + PEM

Inclusion criteria are (1) participants age 18 years or older, (2) PEM indicated by the DSQ-PEM with thresholds following the National Institute of Neurological Disorders and Stroke Common Data Elements PEM Working Group (at least one of the items 1-5 with ≥2 on both the severity and frequency scale) [25], (3) Bell CFIDS disability scale [26] score in the range of 30 to 70, (4) diagnosed chronic inflammatory condition such as rheumatoid arthritis or Sjögren's syndrome. Exclusion criteria are (1) diagnosed PCS, (2) other diagnosed conditions associated with fatigue and symptoms that are not

classified as chronic inflammatory conditions (ICD-10 codes D89.9, M30-M36, K50-K52, K75.4) such as ME/CFS and those mentioned above for group 1, (3) COVID-19 post-vaccination syndrome, (4) poorly controlled chronic conditions with predominant fatigue symptoms (see detailed criteria as described for group 1), (5) inability to provide informed consent, (6) limited German language proficiency, and (7) insufficient technical regarding wearable or mobile phone handling.

Group 3: Healthy Controls

In the third group, we intend to include control probands (1) aged 18 years or older, and (2) Bell CFIDS disability scale 80 or higher. Exclusion criteria are (1) PEM indicated through DSQ-PEM with thresholds following the National Institute of Neurological Disorders and Stroke Common Data Elements PEM Working Group (none of the items 1-5 are rated with a frequency and severity 2 or higher in intensity and frequency), as well as (2) conditions that are also associated with considerable fatigue and PEM (including PCS, post-vaccination syndrome, and those defined above), as well as (3) inability to provide informed consent, (4) limited German language proficiency, and (5) insufficient technical regarding wearable or mobile phone handling.

Sample Size

Due to the exploratory nature of this observational study and the absence of established effect sizes for PEM triggers in literature, formal power calculations are not feasible. While previous studies analyzed several PEM triggers [18], key questions regarding exact triggers, their effect magnitude, and temporal relationships (lag times) remain largely unanswered. Therefore, we chose a pragmatic approach to sample size determination.

A target of 100 participants per group was chosen based on several considerations: First, this sample size provides sufficient data points for machine learning algorithms to identify patterns and develop predictive models, accounting for the expected variability in trigger types and individual responses. Second, with continuous monitoring over up to six months, we anticipate capturing multiple crash

events per participant, yielding a robust dataset for analysis. Third, this sample size allows for subgroup analyses while maintaining reasonable precision for exploratory findings. Finally, recruitment of n=100 participants per group is feasible within the planned timeframe and given available study resources.

Measures

Measures include (a) sociodemographic and disease-related information at baseline, (b) sampling of biomaterials at baseline, (c) patient-reported outcome measures (PROMs) at multiple timepoints, (d) continuous activity monitoring through wearables, and (e) daily PEM symptom monitoring.

(a) Sociodemographic and clinical baseline data

Following informed consent (baseline time point, T0), participants complete a questionnaire administered through evasys (evasys GmbH, Lüneburg, Germany). Sociodemographic data collected include age, gender, characterization of place of residence, educational level, employment status, occupation, living situation, and history of migration.

Clinical baseline data encompasses medical history including current comorbidities and medications. For PCS patients, detailed COVID-19 infection history is collected including date of initial infection, severity of acute illness, number of infections, and vaccination status. All participants report current symptom severity over the past seven days for 33 different symptoms on a scale from 0 (no symptoms) to 10 (most severe symptoms), covering respiratory, cardiovascular, neurological, gastrointestinal, and other organ systems. The selection of symptoms was based on the WHO definition and the UK NICE guidelines on PCS [9, 27]. Anthropometric measurements including height and weight are recorded.

(b) Biomaterial sampling at baseline

At baseline visits, blood samples are collected from participants. The blood collection includes oneserum separator and one Ethylenediaminetetraacetic acid (EDTA) tube. Collected samples undergo

laboratory analysis for a panel of inflammatory markers including interleukin-6 (IL-6), C-reactive protein (CRP), macrophage inflammatory proteins (MIP), and additional chemokines and cytokines relevant to inflammatory processes in both PCS and rheumatic conditions. The further laboratory includes a differential blood count and a basic metabolic panel. To characterize immune responses following infection, SARS-CoV-2 antibody levels will be determined.

Following completion of the planned analyses, any remaining biomaterial is processed and stored in a central biobank according to standardized biobanking protocols. All sample collection and processing procedures follow standardized protocols to ensure consistency across both study sites and sample integrity for downstream analyses.

(c) Patient-reported outcome measures (PROMs)

As shown in Figure 2, standardized instruments are administered at baseline (T0), 3 months (T1), and 6 months (T2). Fatigue and post-exertional malaise are assessed through multiple instruments. The Fatigue Assessment Scale (FAS) comprises 10 items measuring physical and mental fatigue with responses ranging from "never" to "always". The FAS is a reliable and valid instrument for measuring fatigue with good internal consistency and a Cronbach's alpha of 0.82 [28] and has been used in several studies in PCS patients, including studies conducted in Germany [29, 30]. The DSQ-PEM evaluates both frequency and severity of five core PEM symptoms over the past six months, with additional questions about recovery time and activity avoidance [31]. The Bell Disability Scale provides a single score from 0 to 100 indicating functional capacity, while the Canadian Consensus Criteria for ME/CFS assesses symptom clusters across seven domains including post-exertional neuroimmune exhaustion, sleep dysfunction, pain, and neurological/cognitive manifestations [26].

Mental health and quality of life assessments include the (PHQ-9 for depression screening [32], with nine items rated on frequency over the past two weeks, and the Generalized Anxiety Disorder-7 (GAD-7) scale measuring anxiety symptoms. The PHQ-9 is a widely used screening tool for depressive

symptoms and demonstrates a strong reliability and external validity. The internal consistency of the scale is reflected by Cronbach's alpha of 0.82 [32]. Similarly, the commonly used GAD-7 has an internal consistency with Cronbach's alpha of 0.92 [33]. Health-related quality of life is captured through the EuroQol 5-Dimension 5-Level (EQ-5D-5L), assessing mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, plus a visual analog scale for overall health status. The EQ-5D-5L is often used in long COVID research and measures health related quality of life and demonstrates a strong reliability and validity with a good internal consistency and a Cronbach's alpha of 0.75 [34]. The Perceived Stress Scale (PSS-10) evaluates perceived stress over the past month through 10 items measuring feelings of control and coping ability. It also is a reliable and valid tool for measuring perceived stress, demonstrating good internal consistency with a Cronbach's alpha of 0.78 [35].

Functional capacity and social factors are assessed through several instruments. The IMET (Index zur Messung von Einschränkungen der Teilhabe/ Participation Restriction Measurement Index) measures participation restrictions across nine life domains on a scale from 0 (no impairment) to 10 (no activity possible). It has been used by various authors to evaluate social impairment during the COVID-19 pandemic [36–39]. The Work Ability Index short version evaluates current work capacity compared to lifetime best, incorporating both physical and mental work demands and is a validated and frequently used instrument for assessing work ability in PCS patients [29, 40]. Social support is measured through the ENRICHD Social Support Inventory (ESSI) with five items assessing emotional and instrumental support availability. It is also used in post COVID research [9, 41] and has a good internal consistency with a Cronbach's alpha of 0.88 [42]. The Resilience Scale-13 (RS-13) captures psychological resilience through 13 items rated from 1 (disagree) to 7 (agree) and is a reliable and valid tool to measure resilience with a Cronbach's alpha of 0.92 [43]. The UCLA-Loneliness-Scale in the 3-item short form (UCLA-LS) measures the loneliness of the participants [44].

Disease-specific assessments include the EFK (Essener Fragebogen zur Krankheitsverarbeitung/ Essen questionnaire on coping with illness), a 45-item questionnaire measuring various coping strategies

from problem-focused to emotion-focused approaches [45]. The Qualtiy of Life in Neurological Disorders (Neuro-QoL) scales assess sleep disturbances (8 items), stigma (7 items), and cognitive function (8 items) specific to neurological conditions. Internal consistency (Cronbach's α) ranges from 0.85 to 0.97 [46]. The MYMOP-2 (Measure Yourself Medical Outcome Profile) allows participants to identify and rate their most bothersome symptoms and activity limitations on a scale from 0 to 6 [47]. Participants who identify as female are also asked about their menstrual cycle. The Period QoL assesses the quality of life during the period, while the Premenstrual Syndrome (PMS) questionnaire aims to assess the physical and mental complaints during the entire cycle [48, 49].

Additional assessments include the Freiburg Physical Activity Questionnaire documenting various physical activities over the past four weeks, including occupational, transportation, household, and sports activities [50]. Substance use is screened using an adapted version of the WHO ASSIST (Alcohol, Smoking and Substance Involvement Screening Test), covering lifetime use and recent consumption patterns of various substances including alcohol, tobacco, cannabis, and other drugs [51, 52].

Insights into the participants' lifestyle are obtained through the Simple Lifestyle Indicator Questionnaire (SLIQ). This instrument collects data on diet, daily physical activity, alcohol consumption, and everyday stress [53].

(d) Continuous wearable monitoring

Continuous physiological and activity monitoring is conducted using Apple watch SE or another compatible Apple Watch (Apple Inc., Cupertino, CA). Probands can use their own Apple watches or iPhones (iPhone 11 and higher) in case they own such devices and these are compatible with study procedures, otherwise they will be provided with technical devices upon enrollment. The watches will be worn for 90 (non-PEM controls) or 180 (PEM groups) days. Participants are instructed to wear the device for at least 18 hours a day, removing it only for recharging or activities that might damage the device or when there is a risk of injury, such as contact sports.

Apple watches SE or comparable Apple watch models continuously capture health metrics including cardiovascular parameters, like heart rate with maximum, minimum, and resting values in beats per minute, heart rate variability, walking heart rate average, and respiratory rate in breaths per minute. Physical activity and energy expenditure data encompass step count, activity energy and basal energy in kilocalories, distance in meters, flights climbed, exercise minutes, stand hours, and specific activity types (walking, cycling, stationary, or movable activities) measured through motion sensors with speed in meters. Sleep patterns are analyzed through time-based assessment of different sleep phases including REM sleep, deep sleep, core sleep, and awake time, with durations recorded in hours and minutes. For female participants, the Apple watch cycle tracking may request information on menstruation days, flow levels (light, medium, heavy), and associated symptoms.

For this study no actual GPS coordinates or specific location data are collected or stored, ensuring participant privacy while still gathering relevant environmental context. Environmental data is approximated using derived metrics to characterize environmental exposures including location categories (crowded areas, green spaces, urban environments, shopping centers), weather conditions (temperature in degrees Celsius, sky conditions), humidity percentage, and continuous monitoring of noise exposure in decibels to assess potential PEM triggers.

All wearable data is synchronized using a native iOS application.

(e) Daily PEM symptom monitoring

Daily symptom monitoring is conducted through an ultra-short questionnaire assessing two key parameters: whether they experienced a crash in the past 24 hours and their current energy level. Each day, participants receive an automated text message to their mobile phones prompting them to respond to these two questions. This minimally burdensome approach ensures high response rates while capturing essential daily fluctuations in PEM symptoms.

If participants report a crash or significant decline in energy levels, they are automatically directed to an extended survey. This adaptive assessment captures detailed information about the crash event, including timing of onset, severity rating, duration, and specific symptoms experienced. Participants are also asked to identify potential triggers.

The specific instruments for both the daily screening questions and extended crash assessments are developed through an exploratory qualitative interview study with PEM patients. This formative research phase involves qualitative interviews with individuals experiencing PEM across different conditions to ensure the assessment tools capture the most relevant aspects of crash experiences, use appropriate terminology that resonates with patients' lived experiences, and maintain feasibility for completion during periods of severe symptoms. The qualitative findings will be incorporated into the final questionnaire design, including response scales, symptom descriptors, and trigger categories. The interview guide is provided as supplementary material.

	Screening	Group allocation and Baseline	Follow up & Close Out	
	$-t_1$	t_0	t_1 3 months	t_2 6 months
ELIGIBILITY				
Age	X			
Informed consent		X		
Qualifying Diagnosis for Inclusion	X	X		
DSQ-PEM	X	X		X
Bell Score	X	X		X
Pre-existing conditions	X	X		
IT Literacy	X	X		
Group allocation		X		
BIOLOGICAL MEASURES				
Height & Weight		H&W		W
Blood sampling		X		
SOCIAL AND CLINICAL INFORMATION				
Sociodemographics		X		
Medical history and medications		X		X
COVID-19 infection history		X		
Symptom Severity		X	X	X
MEASURES AND ASSESMENT				
FAS		X	X	X
Canadian Criteria for CFS/ME		X		X
Freiburg Physical Activity Questionnaire		X		X
PHQ-9		X	X	X
GAD-7		X	X	X
EQ-5D		X	X	X
Neuro-QOL		X	X	X
WHO ASSIST		X	X	X
IMET		X	X	X
EFK		X	X	X
RS-13		X	X	X
PSS-10		X	X	X
WAI		X		X
ESSI		X	X	X
Period QOL		X	X	X
MYMOP 2		X	X	X
CONTINUOUS DATA				
Heart rate		◆—	—	—◆
Breathing rate		◆—	—	—◆
Sleeping pattern		◆—	—	—◆
Hourly steps & Climbed Flights		◆—	—	—◆
Type and Duration of Activity		◆—	—	—◆
Mode of transportation		◆—	—	—◆
Location type		◆—	—	—◆
Weather		◆—	—	—◆
Noise exposure		◆—	—	—◆
REGULAR BRIEF CHECK-IN				
Surveys on symptom loads		◆—	—	—◆
Crash experience		◆—	—	—◆
Quality of life		◆—	—	—◆

Figure 2: Schedule enrollment, interventions, and assessments.

Monitoring

Our monitoring strategy aims to ensure consistently high data quality and encompasses pre-study feasibility testing, continuous data quality assessment, and proactive participant engagement. Prior to launch, we conduct backend system testing, verify data synchronization between Apple Watch SE or another compatible Apple Watch devices and the corresponding platform, and validate user interfaces to ensure reliable performance.

Other studies using Apple Watches have shown acceptable validity for our measurements. For example, heart rate monitoring showed clinically acceptable accuracy during rest and exercise [54] and activity tracking yielded reasonable validity with mean absolute percentage errors of 18% and intraclass correlation coefficients of 0.69-0.95 for step counting [55]. Sleep tracking achieves 50-86% sensitivity for sleep stage discrimination and 88% agreement for basic sleep-wake categorization [56, 57].

Automated data quality checks identify missing data, irregular wear patterns, and technical anomalies. Participants showing >12 hours of non-wear time or >48 hours without device synchronization receive proactive outreach for troubleshooting and support. Daily PEM symptom assessment response rates are monitored in real-time, with automated reminders sent to non-responders after 12 hours. Weekly data completeness reports enable early identification of participants requiring additional support. All monitoring activities are documented to inform protocol modifications and establish best practices for future digital health studies in PEM populations.

Patient involvement

The study was designed and is continuously supervised by our established Citizen's Advisory Board ("Bürger*innenbeirat") at the Department of General Practice at UMG, which consists of 10-15 individuals with diverse sociodemographic backgrounds, the majority being affected by PCS or other chronic conditions. This board, founded in 2021, provides input on study materials, recruitment strategies, selection of patient-reported outcome measures, and reviews the corresponding data recording application. Throughout the study, preliminary results are discussed with the board.

Research Ethics and Data Safety

This study is conducted in accordance with the ethical principles of the Declaration of Helsinki and has been approved by the ethics committee of Hannover Medical School (11475_BO_S_2024,

13/03/2025). The observational nature of the study presents minimal risk to participants, with data security being the primary consideration. All participant data are managed in pseudonymized form, with direct identifiers replaced by study-specific identification numbers (ID) immediately after enrollment. The key linking these codes to participant identities is stored in encrypted, password-protected files accessible only to authorized study personnel at respective enrollment sites, with strict access controls and audit trails implemented. Following the privacy by design principle, unique device identifiers such as International Mobile Equipment Identity (IMEI) numbers of smartwatches or phones are deliberately excluded from research databases to prevent potential re-identification of participants. All data access is governed by role-based permission systems, ensuring that investigators only have access to the minimum data set necessary for their specific research tasks. Participants receive comprehensive information regarding data protection measures during the informed consent process, including transparent explanation of data handling procedures, potential risks, and their rights regarding data withdrawal. All procedures comply with the General Data Protection Regulation (GDPR) and applicable national data protection laws.

Data Processing and Storage

The data collection streams include paper-based questionnaires at baseline (T0), after 3 (T1) and 6 (T2) months, and specimen sampling completed at enrollment, daily monitoring of PEM symptoms via a dedicated app, and continuous wearable metrics from the Apple watch SE or another compatible Apple watch.

Baseline questionnaires are digitized through optical character recognition (OCR) technology following completion at enrollment visits using evasys (evasys GmbH, Lüneburg, Germany). The scanned documents undergo manual verification by study personnel to ensure accuracy of the digitized data prior to integration into the central database. Completed paper questionnaires are archived for 10 years.

Blood samples collected at baseline are processed at a centralized, certified laboratory facility at Hannover, Germany (Laboratory Stille/Jablonka Hannover and MHH). Upon arrival in the laboratory, samples are centrifuged, and plasma is isolated and frozen to -20°C until batch analysis. All samples are analyzed using a COBAS ® system (Roche Diagnostics International AG, Rotkreuz, Switzerland) according to the manufacturer's recommendations, with surplus material biobanked for future research use following participant consent.

Electronic check-ins for specific events are recorded via the study app. The application operates on ISO 27001:2013 certified servers located at GWDG in Germany. Security measures include end-to-end secure sockets layer (SSL) encryption for all data transmission and storage.

We collect data from Apple smartwatches through a native iOS application developed specifically for this study. Study smartwatches are equipped with an eSIM providing independent connectivity, enabling continuous data collection without requiring a connected iPhone nearby. This autonomous connectivity ensures uninterrupted data capture even when participants are not carrying their phones. All wearable data is transmitted via secure virtual private network (VPN) connections to a protected PostgreSQL database at GWDG, with personal identifiers encrypted and linked only to pseudonymized user IDs to maintain participant privacy while enabling data integration across sources.

Temporal alignment algorithms synchronize time-stamped wearable measurements with participant-reported crash events, enabling precise analysis of physiological changes preceding, during, and following PEM episodes. All details regarding data collection, processing, storage, access control, and long-term preservation are comprehensively documented in the Data Management Plan (DMP), which is available as an appendix to this protocol.

Data Analysis

The primary objective is to identify triggers for PEM episodes and to develop algorithms for wearable based, real time prediction of PEM crashes. We will model time to event using both semi-parametric and machine learning survival approaches that handle repeated events and irregular observation intervals.

Data preparation

The dataset includes mainly ordinal or numerical variables derived from various sources (see Data Processing and Storage). All timestamps will be synchronized and converted to the "Europe/Berlin" time zone. Depending on the original sampling resolution and clinical relevance, time-dependent variables will be resampled into non-overlapping time windows. As no prior information about the signal quality or temporal dynamics of the incoming data is available at the time of protocol development, final resolution choices will be based on visual inspection, data completeness, clinical relevance and preliminary analyses. Within each fixed window, we will store summary statistics such as mean, median, standard deviation, minimum, and maximum. For hypothesized delayed effects, lag-features (e.g. 24h, 48h, 72h) will be generated. Change-point detection using the PELT algorithm [58] will be applied to selected physiological features and encoded as binary indicators (change occurred vs. not) or as time since last change.

If necessary to reduce multicollinearity, only one representative variable from a cluster of highly correlated variables ($|\rho| > 0.9$) will be retained in the final dataset. This representative variable will be selected based on clinical interpretability, temporal relevance (e.g., shorter lags preferred), or signal stability across participants. The correlation threshold ρ may be adjusted based on the overall correlation structure and its impact on model robustness. Additionally, variables with near-zero variance, excessive missingness (>50%), poor signal stability across individuals, or extremely unbalanced categories may be excluded. Accordingly, samples with >50% missingness may be

removed, with reasons documented. The 50% threshold may be refined based on exploratory missingness analysis, including missingness patterns, clinical relevance, and imputation feasibility. Missing values with minimal missingness (<5%) in static baseline variables and low-frequency questionnaire items will be imputed applying median or mode. For time-varying covariates and longitudinal measurements multivariate imputation by chained equations (20 imputations) will address sporadic missingness (<15 %). Variables with 15–50% missingness will generally be included in multivariate imputation, unless excluded due to poor imputation quality or low clinical relevance. Physiologically implausible values will be winsorized at the 0.5th/99.5th percentiles. Continuous features will be z-standardized, and categorical variables will be one-hot encoded.

Pre-processing sensitivity checks

To evaluate robustness, all modelling steps will be repeated under three alternative preprocessing schemes: (1) single imputation of missing longitudinal values using last-observation-carried-forward (LOCF), (2) omitting winsorization and using raw values, and (3) applying min–max scaling instead of z-score standardization. Significant discrepancies in key performance metrics will be reported to assess the impact of preprocessing choices on model robustness.

Group Comparisons

We will perform predefined exploratory group comparisons to assess whether the collected data contain actionable early-warning signals for PEM events. First, each predefined subgroup (PCS-associated PEM, chronic inflammatory disease-associated PEM) will be analyzed separately. Within each subgroup, we will compare exposure windows preceding events to control windows without events, evaluating whether the algorithm’s predictions significantly exceed chance-level performance and yield clinically relevant lead-times.

Survival modelling

The principal model will be the Cox proportional hazards model [59, 60], using time since study entry (i.e. time-on-study) as the analysis time scale. If necessary, time-dependent covariates will be incorporated by structuring the data in start–stop format. To reduce overfitting and improve model generalizability, penalized regression methods such as elastic-net regularization ($\alpha = 0.5$) may be considered. The necessity and type of penalization will be decided upon preliminary analyses. For unpenalized Cox models with multiple observations per subject, robust sandwich variance estimators will be applied to account for intra-subject correlation. Model assumptions will be evaluated using residual diagnostics (e.g. Schoenfeld and Martingale plots). Non-linear relationships will be addressed using restricted cubic splines with three knots where appropriate. Additionally, an additive hazards model [61, 62] will be fitted as a robustness check, providing an alternative framework that does not rely on proportional hazards assumptions.

To further explore non-linear effects and assess the robustness of findings, conditional inference survival trees (ctree) [63] and random survival forests (RSF) [64] will be implemented as complementary approaches. Key RSF hyperparameters will be tuned, including the number of trees, the number of candidate variables considered at each split, and the minimum terminal node size. For additional benchmark analyses, strongly regularized neural network-based survival models (e.g. DeepSurv [65], DeepHit [66]) will be explored. Neural networks will be evaluated cautiously as non-primary, exploratory methods. Model complexity will be minimized by strongly restricting hidden layers and neurons per layer and dropout regularization. Hyperparameters such as learning rate, dropout rate, and batch size will be optimized via Optuna [67] with early stopping, limiting the search to a maximum of 50 trials. Search ranges will be constrained based on prior benchmarks and pilot analyses to mitigate overfitting.

Model validation and performance metrics

Internal model validation will follow a two-tiered strategy. First, nested cross-validation will be used for hyperparameter tuning. For models that handle temporally ordered data (e.g. penalized Cox, RSF), each inner and outer loop adopts a rolling-origin, grouped-fold resampling scheme to preserve temporal dependencies and prevent information leakage. Neural networks will rely on early stopping against a temporally held-out validation fold within each training split. Second, a temporal hold-out will be used to retrain every final model on the first 90 days of follow-up and evaluated on the subsequent 90 days to quantify generalizability. Model performance will be evaluated using time-dependent discrimination metrics (Harrell's C-index and area under the curve (AUC) at clinically relevant time points) and calibration metrics (integrated Brier score, calibration plots). Uncertainty will be quantified as follows: 1,000 bootstrap resamples of the temporal test set for Cox and additive models, out-of-bag predictions repeated with 20 different random seeds for RSF (reporting median ± interquartile range (IQR) of metrics) and 200 bootstraps for penalized Cox. For explorative neural networks mean ± standard deviation across five rolling-origin folds will be calculated.

Feature selection and stability assessment

Feature selection in penalized models will be based on non-zero coefficients after regularization. In unpenalized models, selection will rely on effect sizes, p-values, and model selection criteria such as the Akaike Information Criterion (AIC) [68]. In the random survival forest (RSF), variable importance will be estimated via permutation. For neural network models, feature contributions will be assessed using SHAP (SHapley Additive exPlanations) values [69]. To evaluate feature selection stability, we will count how frequently each feature is selected or ranks among the top predictors across repeated resampling. Features retained in ≥50% of resamples will be considered robust.

Subgroup Analysis

Models will be re-estimated in the predefined subgroups (PCS associated PEM, chronic inflammatory disease associated PEM) to explore potential heterogeneity in associations and outcomes. Where appropriate, analyses may also be stratified by additional clinically relevant characteristics such as age, sex, or disease type or severity. Interaction terms may also be included in the main models to formally test for heterogeneity. These analyses are considered exploratory and hypothesis-generating.

Case-crossover analysis

To complement between-subject models, we will perform a bidirectional case-crossover analysis in which each participant serves as his or her own control. For each event, we will define exposure windows and compare them to control windows free of events within ±7 days, using conditional logistic regression [70]. This design inherently adjusts for stable, unmeasured confounders and will help to distinguish acute triggers from baseline risk factors.

Sensitivity analyses

Sensitivity analyses will explore: (1) different lag-window definitions (e.g. 3 h, 12 h), (2) alternative imputation methods (missForest [71]), (3) exclusion of outliers instead of winsorisation, and (iv) altered RSF hyperparameter settings. Results will be tabulated alongside primary estimates.

Software

Data preparation and analysis will be performed in various programs, such as SPSS Version 30 (IBM Corp., Armonk NY, USA) as well as up-to-date versions of Python and R, depending on the research context. Core R packages include survival [60, 72], timereg [62], mlr3proba [73], mice [74], partykit [75], and randomForestSRC [64]. Preprocessing relies on lubridate [76] , slider [77], changepoint [58], and DescTools [78]. Neural networks (DeepSurv, DeepHit) will be implemented in Python via pycox

[79], interpreted with SHAP [69] and tuned with Optuna [67]. Final software versions, code and workflow will be archived (e.g. GitHub/Zenodo) to ensure reproducibility.

Discussion

The U-WaTCH study represents an exploratory observational investigation designed to identify crash triggers in participants experiencing PEM. While our current focus is on PCS and rheumatic diseases, the methodological approach may have future applications for other conditions exhibiting PEM symptoms, such as depression and multiple sclerosis.

Strengths and Limitations

A hurdle for interested people may be that only participants with sufficient technical skills and internet access can be included. To address this barrier, we provide clear instructions and ongoing technical to all participants. As devices are provided for study participations, people do not have to own a smartphone and an Apple Watch. Despite these measures, technical challenges remain significant barriers across all age groups.. Key challenges in smartwatch studies include frequent charging requirements, connectivity problems, app synchronization failures, and data quality concerns [80]. Our monitoring strategy specifically addresses these challenges through proactive technical support and automated data quality checks. Beukenhorst et al. [81] reported good engagement over a 3-month period with participants wearing devices on 75% of days or more, though participation declined in the final weeks. This pattern informed our decision to implement different observation periods for control participants (3 months) versus PEM groups (6 months), balancing data collection needs with expected adherence patterns.

In spite of possible technical challenges, the largely digital nature of the U-WaTCH study design is also a strength, as it makes our approach easily accessible to persons that are less mobile (e.g. those affected by PCS, rheumatoid diseases, and fatigue) and scalable.

Including PCS patients, rheumatoid disease patients, and healthy controls allows a precise identification of cross-group differences und distinguishes physiological changes between PCS associated PEM and PEM in patients with rheumatoid disease. This comparison enables identification of disease-specific patterns, potentially disclosing targeted treatment options.

Another strength of our study design is the comprehensive and structured data collection, creating a benchmark for categorized PCS patient data combining wearable data, biomarkers, and PROMs. In this context, a possible limitation is that wearables may produce inaccurate data due to technical issues or irregular wearing. Our study design addresses these challenges common to wearable-based longitudinal research by immediate feedback to those study participants with gaps in data submission over a specific time threshold. This may be particularly important in older study participants, as Chandrasekaran et al. [82] found that approximately 30% of US adults use wearable healthcare devices, but that utilization decreases among higher age groups.

Loss to follow-up represents a particular concern for longer observation periods. Mahalingaiah et al. [83] reported only 35% response rates to follow-up surveys at 6 months in the Apple Women's Health Study. Our multi-modal engagement strategy combines automated reminders, personalized outreach and ultra-short questionnaires to mitigate dropout risks. Given the daily impact of PEM on participants' lives, we anticipate that the potential insights gained from participation may provide intrinsic motivation beyond what is typically seen in healthy population studies, though this remains to be empirically tested.

Abbreviations

AIC - Akaike Information Criterion

AUC - Area Under The Curve

CFIDS - Chronic Fatigue and Immune Dysfunction Syndrome

CIAS - Computerized Intervention Authoring System (cias.app)

CRP - C-reactive protein

DMP - Data Management Plan

DSQ-PEM - DePaul Symptom Questionnaire-PEM

EDTA - Ethylenediaminetetraacetic acid

EFK - Essener Fragebogen zur Krankheitsverarbeitung

EQ-5D-5L - EuroQol 5-Dimension 5-Level

ESSI - ENRICHD Social Support Inventory

FAS - Fatigue Assessment Scale

GAD-7 - Generalized Anxiety Disorder-7

GDPR - General Data Protection Regulation

GPS - Global Positioning System

ICD-10 - International Classification of Diseases, 10th Revision

ID - Identification

IL-6 - Interleukin-6IMET - Index zur Messung von Einschränkungen der Teilhabe

IQR - Interquartile Range

LOCF - Last-observation-carried-forward

ME/CFS - Myalgic Encephalomyelitis/Chronic Fatigue Syndrome

MIP - Macrophage inflammatory proteins

MYMOP-2 - Measure Yourself Medical Outcome Profile

Neuro-QoL - Qualtiy of Life in Neurological Disorders

NYHA - New York Heart Association

OCR - Optical Character Recognition

OSAS - Obstructive Sleep Apnea Syndrome

PCS - Post-COVID Syndrome

PEM - Post-exertional malaise

PELT - Pruned Exact Linear Time

PHQ-9 - Patient Health Questionnaire-9

PMS - Premenstrual Syndrome

PROMs - Patient-reported outcome measures

PSS-10 - Perceived Stress Scale

RS-13 - Resilience Scale-13

RSF - Random survival forests

SARS-CoV-2 - Severe Acute Respiratory Syndrome Coronavirus 2

SHAP - SHapley Additive exPlanations

SLIQ - Simple Lifestyle Indicator Questionnaire

SSL - Secure Sockets Layer

TSH - Thyroid-stimulating hormone

UCLA-LS – UCLA Loneliness Scale

VPN - Virtual Private Network

WHO - World Health Organization

WHO ASSIST - WHO Alcohol, Smoking and Substance Involvement Screening Test

Declarations

Ethics approval and consent to participate

The study is registered at the German registry for clinical trials (DRKS00036301, 27/03/2025) and has been approved by the research ethics board of the Hannover Medical School (11475_BO_S_2024, 13/03/2025).

All participants will receive information on study procedures, participant information and data protection information prior to inclusion. Study participation is voluntary at any time, and participants have the right to withdraw consent without disclosure of reasons for withdrawal. Members of the study team are available for questions at enrollment and thereafter.

Consent for publication

Not applicable.

Relevant guidelines and regulations

This study will be conducted in accordance with Good Clinical Practice (GCP) guidelines.

Availability of data and materials

Data is not publicly available due to a decision of the responsible research ethics board. Data can be obtained upon reasonable request and within a data sharing agreement from the corresponding author.

Competing interest

The authors declare that they have no competing interests.

Funding

The U-WaTCH Study is funded by the German Federal Ministry of Health (ZMII2-2524FSB74B). The funder provided financial resources for the project but had no role in the design of this study and will not have any role during its execution, analyses, interpretation of the data, or decision to submit results.

Authors' Contributions

Study conceptualization and design, Methodology development, Funding acquisition: CH, SS, AD-J, GMN, FK, FM, DS

Clinical recruitment and patient management: VL, TR, LMT, FM, AD-J, RR, TS

Technical infrastructure and platform development: PD, PW

Machine learning algorithm development: SM, FK, PD, PW

Statistical analysis planning: SM, FK

Data management and curation: PD, LMT, TR, VL

Laboratory analysis and biobanking: GMN, AD-J, CH, RR

Patient and public involvement coordination: TR, VL, LMT, TS

Software development and maintenance: PD, PW

Quality assurance and monitoring: PD, FM, CH

Writing – original draft preparation: VL, TR, LMT, FM, CH, EN

Writing – review and editing: All Authors

Supervision: GMN, EH

All authors read and approved the final manuscript.

Acknowledgements

We acknowledge all persons involved in the Citizen's Advisory Board ("Bürger*innenbeirat Allgemeinmedizin") of UMG which helped us to encompass the patient perspective at all levels of study protocol development.

References

1. National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Global Health; Board on Health Sciences Policy; Committee on Examining the Working Definition for Long COVID. A Long COVID Definition: A Chronic, Systemic Disease State with Profound Consequences. Washington (DC): National Academies Press (US); 2024.

2. Hossain MM, Das J, Rahman F, Nesa F, Hossain P, Islam AMK, et al. Living with "long COVID": A systematic review and meta-synthesis of qualitative evidence. PloS One. 2023;18:e0281884.

3. McNabb KC, Bergman AJ, Smith-Wright R, Seltzer J, Slone SE, Tomiwa T, et al. "It was almost like it's set up for people to fail" A qualitative analysis of experiences and unmet supportive needs of people with Long COVID. BMC Public Health. 2023;23:2131.

4. Schmachtenberg T, Königs G, Roder S, Müller F, Müllenmeister C, Schröder D, et al. How do people with long COVID utilize COVID-19 vaccination and rehabilitation services and what are their experiences with these services? results of a qualitative study with 48 participants from Germany. BMC Public Health. 2024;24.

5. Schröder D, Heinemann S, Heesen G, Hummers E, Schmachtenberg T, Dopfer-Jablonka A, et al. Association of long COVID with health-related Quality of Life and Social Participation in Germany: Finding from an online-based cross-sectional survey. Heliyon. 2024;10:e26130.

6. Hastie CE, Lowe DJ, McAuley A, Mills NL, Winter AJ, Black C, et al. True prevalence of long-COVID in a nationwide, population cohort study. Nat Commun. 2023;14:7892.

7. Post COVID-19 condition (long COVID). https://www.who.int/news-room/fact-sheets/detail/post-covid-19-condition-(long-covid). Accessed 19 May 2025.

8. Goldenberg DL. Applying Lessons From Rheumatology to Better Understand Long COVID. Arthritis Care Res. 2024;76:49–56.

9. Soriano JB, Murthy S, Marshall JC, Relan P, Diaz JV. A clinical case definition of post-COVID-19 condition by a Delphi consensus. Lancet Infect Dis. 2022;22:e102–7.

10. Vernon SD, Zheng T, Do H, Marconi VC, Jason LA, Singer NG, et al. Incidence and Prevalence of Post-COVID-19 Myalgic Encephalomyelitis: A Report from the Observational RECOVER-Adult Study. J Gen Intern Med. 2025;40:1085–94.

11. Tokumasu K, Honda H, Sunada N, Sakurada Y, Matsuda Y, Yamamoto K, et al. Clinical Characteristics of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) Diagnosed in Patients with Long COVID. Medicina (Mex). 2022;58:850.

12. Hartle M, Bateman L, Vernon SD. Dissecting the nature of post-exertional malaise. Fatigue Biomed Health Behav. 2021;9:33–44.

13. Renz-Polster H, Scheibenbogen C. Post-COVID-Syndrom mit Fatigue und Belastungsintoleranz: Myalgische Enzephalomyelitis bzw. Chronisches Fatigue-Syndrom. Inn Med. 2022;63:830–9.

14. Kim L, Kedor C, Buttgereit F, Heidecke H, Schaumburg D, Scheibenbogen C. Characterizing Sjögren-Associated Fatigue: A Distinct Phenotype from ME/CFS. J Clin Med. 2023;12:4994.

15. Vernon SD, Hartle M, Sullivan K, Bell J, Abbaszadeh S, Unutmaz D, et al. Post-exertional malaise among people with long COVID compared to myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). Work. 2023;74:1179–86.

16. Greenwood DC, Mansoubi M, Bakerly ND, Bhatia A, Collett J, Davies HE, et al. Physical, cognitive, and social triggers of symptom fluctuations in people living with long COVID: an intensive longitudinal cohort study. Lancet Reg Health - Eur. 2024;46:101082.

17. Klebek L, Sunnquist M, Jason LA. Differentiating Post-Polio Syndrome from Myalgic Encephalomyelitis and Chronic Fatigue Syndrome. Fatigue Biomed Health Behav. 2019;7:196–206.

18. Pouliopoulou DV, Hawthorne M, MacDermid JC, Billias N, Miller E, Quinn K, et al. Prevalence and Impact of Postexertional Malaise on Recovery in Adults With Post-COVID-19 Condition: A Systematic Review With Meta-analysis. Arch Phys Med Rehabil. 2025;:S0003-9993(25)00501-5.

19. Suche | Landesamt für Statistik Niedersachsen. https://www.statistik.niedersachsen.de/live/search.php. Accessed 30 Jun 2025.

20. Marcu G, Ondersma SJ, Spiller AN, Broderick BM, Kadri R, Buis LR. Barriers and Considerations in the Design and Implementation of Digital Behavioral Interventions: Qualitative Analysis. J Med Internet Res. 2022;24:e34301.

21. Mikuteit M, Heinemann S, Roder S, Niewolik J, Schröder D, Vahldiek K, et al. Long-term Consequences of COVID-19 and the Pandemic: Protocol for a Web-Based, Longitudinal Observational Study (DEFEAT). JMIR Res Protoc. 2022;11:e38718.

22. Organization WH. A clinical case definition of post COVID-19 condition by a Delphi consensus, 6 October 2021. 2021.

23. Guidance for Core PEM Assessment (DePaul Symptom Questionnaire Subscale).

24. Sunnquist M, Lazarus S, Jason LA. The development of a short form of the DePaul Symptom Questionnaire. Rehabil Psychol. 2019;64:453–62.

25. NINDS CDE Project Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) Post Exertional Malaise Subgroup. 2017.

26. Bell DS. The Doctor’s Guide to Chronic Fatigue Syndrome: Understanding, Treating and Living with CFIDS.

27. COVID-19 rapid guideline: managing the long-term effects of COVID-19. 2024.

28. Hendriks C, Drent M, Elfferich M, De Vries J. The Fatigue Assessment Scale: quality and availability in sarcoidosis and other diseases. Curr Opin Pulm Med. 2018;24:495–503.

29. Kerling A, Beyer S, Dirks M, Scharbau M, Hennemann A-K, Dopfer-Jablonka A, et al. Effects of a randomized-controlled and online-supported physical activity intervention on exercise capacity, fatigue and health related quality of life in patients with post-COVID-19 syndrome. BMC Sports Sci Med Rehabil. 2024;16:33.

30. Beyer S, Haufe S, Dirks M, Scharbau M, Lampe V, Dopfer-Jablonka A, et al. Post-COVID-19 syndrome: Physical capacity, fatigue and quality of life. PloS One. 2023;18:e0292928.

31. Sunnquist M, Lazarus S, Jason LA. The development of a short form of the DePaul Symptom Questionnaire. Rehabil Psychol. 2019;64:453–62.

32. Rahman MA, Dhira TA, Sarker AR, Mehareen J. Validity and reliability of the Patient Health Questionnaire scale (PHQ-9) among university students of Bangladesh. PloS One. 2022;17:e0269634.

33. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. Arch Intern Med. 2006;166:1092.

34. Fernández-de-Las-Peñas C, Rodríguez-Jiménez J, Moro-López-Menchero P, Cancela-Cilleruelo I, Pardo-Hernández A, Hernández-Barrera V, et al. Psychometric properties of the Spanish version of the EuroQol-5D-5L in previously hospitalized COVID-19 survivors with long COVID. Sci Rep. 2022;12:12605.

35. Klein EM, Brähler E, Dreier M, Reinecke L, Müller KW, Schmutzer G, et al. The German version of the Perceived Stress Scale - psychometric characteristics in a representative German community sample. BMC Psychiatry. 2016;16:159.

36. Schröder D, Stölting A, Müllenmeister C, Behrens GMN, Klawitter S, Klawonn F, et al. Improvement in quality of life and cognitive function in Post-COVID syndrome after online occupational therapy: Results from a randomized controlled pilot study. PLOS One. 2025;20:e0312714.

37. Schröder D, Heinemann S, Heesen G, Hummers E, Schmachtenberg T, Dopfer-Jablonka A, et al. Association of long COVID with health-related Quality of Life and Social Participation in Germany: Finding from an online-based cross-sectional survey. Heliyon. 2024;10:e26130.

38. Schröder D, Schmachtenberg T, Heinemann S, Müllenmeister C, Roder S, El-Sayed I, et al. Parenting and Gender as Impact Factors for Social Participation, Quality of Life, and Mental Health in Long COVID. J Prim Care Community Health. 2024;15.

39. Niewolik J, Mikuteit M, Schröder D, Heinemann S, Heesen G, Müller F, et al. Soziale Teilhabe und Hautkrebs während der COVID-19-Pandemie. Dermatol. 2023;74:108–13.

40. Kerksieck P, Ballouz T, Haile SR, Schumacher C, Lacy J, Domenghino A, et al. Post COVID-19 condition, work ability and occupational changes in a population-based cohort. Lancet Reg Health Eur. 2023;31:100671.

41. Schick V, Lieb M, Borho A, Morawa E, Geiser F, Beschoner P, et al. Psychological and Social Factors Associated With Reporting Post-COVID Symptoms Among German Healthcare Workers: A Cross-Sectional Study. J Clin Nurs. 2025. https://doi.org/10.1111/jocn.17608.

42. Vaglio J, Conard M, Poston WS, O'Keefe J, Haddock CK, House J, et al. Testing the performance of the ENRICHD Social Support Instrument in cardiac patients. Health Qual Life Outcomes. 2004;2:24.

43. Rosendahl J, Ebmeyer K, Strauß B, Engert V. [New Normative Values for the German Short Version of the Resilience Scale (RS-13)]. Psychother Psychosom Med Psychol. 2024;74:395–402.

44. Dr. Janosch Schobin (Kompetenznetz Einsamkeit/, Institut für Sozialarbeit und Sozialpädagogik e. V.), Céline Arriagada (Kompetenznetz Einsamkeit/, Institut für Sozialarbeit und Sozialpädagogik e. V.), Martin Gibson-Kunze (Kompetenznetz Einsamkeit/, Institut für Sozialarbeit und Sozialpädagogik e. V.). Einsamkeitsbarometer 2024. März 2024, 1. Auflage.

45. Franke GH, Jagla-Franke M. EFK - Essener Fragebogen zur Krankheitsverarbeitung. 2021.

826 46. Cella D, Lai J-S, Nowinski CJ, Victorson D, Peterman A, Miller D, et al. Neuro-QOL: brief measures
827 of health-related quality of life for clinical research in neurology. Neurology. 2012;78:1860–7.

828 47. Polus BI, Kimpton AJ, Walsh MJ. Use of the measure your medical outcome profile (MYMOP2)
829 and W-BQ12 (Well-Being) outcomes measures to evaluate chiropractic treatment: an observational
830 study. Chiropr Man Ther. 2011;19:7.

831 48. Lancastle D, Kopp Kallner H, Hale G, Wood B, Ashcroft L, Driscoll H. Development of a brief
832 menstrual quality of life measure for women with heavy menstrual bleeding. BMC Womens Health.
833 2023;23:105.

834 49. Ditzen B, Nussbeck F, Drobnjak S, Spörri C, Wüest D, Ehlert U. Validierung eines
835 deutschsprachigen DSM-IV-TR basierten Fragebogens zum prämenstruellen Syndrom. Z Für Klin
836 Psychol Psychother. 2011;40:149–59.

837 50. Messung der Sport- und Bewegungsaktivität — Institut für Sport und Sportwissenschaft.
838 https://www.sport.uni-
839 freiburg.de/de/institut/psychologie/messinstrumente/Messung%20der%20Sport%20und%20Beweg
840 ungsaktivitaet. Accessed 16 Jun 2025.

841 51. Humeniuk R, Ali R, Babor TF, Farrell M, Formigoni ML, Jittiwutikarn J, et al. Validation of the
842 alcohol, smoking and substance involvement screening test (ASSIST). Addiction. 2008;103:1039–47.

843 52. Suchtmedizin AZ für. ASSIST. Arud Zentrum für Suchtmedizin.
844 https://arud.ch/fachbereiche/fachbereiche/psychiatrie-und-psychotherapie/assist. Accessed 16 Jun
845 2025.

846 53. Godwin M, Streight S, Dyachuk E, van den Hooven EC, Ploemacher J, Seguin R, et al. Testing the
847 Simple Lifestyle Indicator Questionnaire: Initial psychometric study. Can Fam Physician Med Fam Can.
848 2008;54:76–7.

849 54. Falter M, Budts W, Goetschalckx K, Cornelissen V, Buys R. Accuracy of Apple Watch
850 Measurements for Heart Rate and Energy Expenditure in Patients With Cardiovascular Disease:
851 Cross-Sectional Study. JMIR MHealth UHealth. 2019;7:e11889.

852 55. Viciana J, Casado-Robles C, Guijarro-Romero S, Mayorga-Vega D. Are Wrist-Worn Activity
853 Trackers and Mobile Applications Valid for Assessing Physical Activity in High School Students?
854 Wearfit Study. J Sports Sci Med. 2022;:356–75.

855 56. Robbins R, Weaver MD, Sullivan JP, Quan SF, Gilmore K, Shaw S, et al. Accuracy of Three
856 Commercial Wearable Devices for Sleep Tracking in Healthy Adults. Sensors. 2024;24:6532.

857 57. Miller DJ, Sargent C, Roach GD. A Validation of Six Wearable Devices for Estimating Sleep, Heart
858 Rate and Heart Rate Variability in Healthy Adults. Sensors. 2022;22:6317.

859 58. Killick R, Eckley IA. **changepoint** : An *R* Package for Changepoint Analysis. J Stat Softw. 2014;58.

860 59. Cox DR. Regression Models and Life-Tables. J R Stat Soc Ser B Stat Methodol. 1972;34:187–202.

861 60. CRAN: survival citation info. https://cran.r-project.org/web/packages/survival/citation.html.
862 Accessed 16 Jun 2025.

863 61. Dynamic Regression Models for Survival Data. New York, NY: Springer New York; 2006.

62. Scheike TH, Zhang M-J. Analyzing Competing Risk Data Using the *R* **timereg** Package. J Stat Softw. 2011;38.

63. Hothorn T, Hornik K, Zeileis A. Unbiased Recursive Partitioning: A Conditional Inference Framework. J Comput Graph Stat. 2006;15:651–74.

64. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. Ann Appl Stat. 2008;2.

65. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. BMC Med Res Methodol. 2018;18:24.

66. Lee C, Zame W, Yoon J, Van Der Schaar M. DeepHit: A Deep Learning Approach to Survival Analysis With Competing Risks. Proc AAAI Conf Artif Intell. 2018;32.

67. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: A Next-generation Hyperparameter Optimization Framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage AK USA: ACM; 2019. p. 2623–31.

68. Cavanaugh JE, Neath AA. The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. WIREs Comput Stat. 2019;11:e1460.

69. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc.; 2017. p. 4768–77.

70. Prentice RL, Breslow NE. Retrospective studies and failure time models. Biometrika. 1978;65:153–8.

71. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012;28:112–8.

72. Therneau TM, until 2009) TL (original S->R port and R maintainer, Elizabeth A, Cynthia C. survival: Survival Analysis. 2024.

73. Sonabend R, Király FJ, Bender A, Bischl B, Lang M. mlr3proba: an R package for machine learning in survival analysis. Bioinformatics. 2021;37:2789–91.

74. Buuren S van, Groothuis-Oudshoorn K, Vink G, Schouten R, Robitzsch A, Rockenschaub P, et al. mice: Multivariate Imputation by Chained Equations. 2025.

75. Hothorn T, Zeileis A. Partykit: a modular toolkit for recursive partytioning in R. J Mach Learn Res. 2015;16:3905–9.

76. Spinu V, Grolemund G, Wickham H, Vaughan D, Lyttle I, Costigan I, et al. lubridate: Make Dealing with Dates a Little Easier. 2024.

77. Vaughan D, Software P, PBC. slider: Sliding Window Functions. 2024.

78. Signorell A, Aho K, Alfons A, Anderegg N, Aragon T, Arachchige C, et al. DescTools: Tools for Descriptive Statistics. 2025.

79. Kvamme H. havakv/pycox. 2025.

80. Triantafyllidis A, Kondylakis H, Katehakis D, Kouroubali A, Alexiadis A, Segkouli S, et al. Smartwatch interventions in healthcare: A systematic review of the literature. Int J Med Inf. 2024;190:105560.

81. Beukenhorst AL, Howells K, Cook L, McBeth J, O'Neill TW, Parkes MJ, et al. Engagement and Participant Experiences With Consumer Smartwatches for Health Research: Longitudinal, Observational Feasibility Study. JMIR MHealth UHealth. 2020;8:e14368.

82. Chandrasekaran R, Katthula V, Moustakas E. Patterns of Use and Key Predictors for the Use of Wearable Health Care Devices by US Adults: Insights from a National Survey. J Med Internet Res. 2020;22:e22443.

83. Mahalingaiah S, Fruh V, Rodriguez E, Konanki SC, Onnela J-P, De Figueiredo Veiga A, et al. Design and methods of the Apple Women's Health Study: a digital longitudinal cohort study. Am J Obstet Gynecol. 2022;226:545.e1-545.e29.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- InterviewguideUWaTCH.docx