BMJ Open

Characterizing Associations of Exercise and Pain Patterns in Endometriosis via Mobile Self-Tracking

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-059280
Article Type:	Original research
Date Submitted by the Author:	27-Nov-2021
Complete List of Authors:	Ensari, Ipek; Columbia University, Data Science Institute Lipsky-Gorman, Sharon; Columbia University, Department of Biomedical Informatics Horan, Emma; Columbia University, Department of Biomedical Informatics Bakken, Suzanne; Columbia University, School of Nursing Elhadad, Noemie; Columbia University, Department of Biomedical Informatics
Keywords:	PAIN MANAGEMENT, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS, PREVENTIVE MEDICINE, EPIDEMIOLOGY, COMPLEMENTARY MEDICINE

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Characterizing Associations of Exercise and Pain Patterns in Endometriosis via Mobile Self-Tracking

Running title: Daily exercise and pain patterns in endometriosis

Ipek Ensari[1], PhD, Sharon Lipsky-Gorman[2], MA, Emma Horan[2], BS, Suzanne R. Bakken[3], PhD, FAAN, Noémie Elhadad[2], PhD

[1]Data Science Institute, Columbia University, [2]Department of Biomedical Informatics, Columbia University Irving Medical Center, [3]Columbia University School of Nursing, New York, NY

Author of Correspondence:
Ipek Ensari
Columbia University
Data Science Institute
475 Riverside Dr, Room 320,
New York, NY 10115
Email:ie2145@columbia.edu

Word count: 4,601

Abstract

Objectives: This study investigates the association of exercise to pain at the day level in endometriosis, an inflammatory chronic pain condition that is currently inadequately managed and could benefit from exercise as a component of its effective management. **Setting**: A participatory research-based smartphone app (Phendo) designed for tracking symptoms and self-management of endometriosis. **Participants:** Study sample included 90,382 days of data from 1,009 Phendo research participants (~85% non-Hispanic white) with self-reported endometriosis living across 38 countries. **Primary Outcome Measures:** 1) Daily pain score that includes its intensity and location, 2) Change in pain score from previous day. **Design:** This was an observational, retrospective study. Pain outcomes were estimated from previous-day exercise and pain symptoms in separate, covariate-adjusted linear mixed-level models. **Results**: The association of previous-day exercise to pain outcomes was moderated by habitual exercise levels, independent of type of endometriosis diagnosis or body mass index (Rate ratio=0.96, 95% CI=0.95, 0.98, p=0.0007 for pain score outcome, B=-0.14, 95%CI=-0.26, -0.016, p=0.026 for pain difference). The habitual exercise level at which previous-day exercise started to be associated with favorable pain outcomes was ~3 times per week. Walking, yoga and stretching type activities were the most frequently reported modalities. **Conclusions:** To accrue the benefits of exercise for adequate endometriosis pain management at the day level, exercise might first need to be developed as a habit. A better understanding of the relationship between exercise behavior and endometriosis pain can be a starting point for identifying optimum points of intervention for informing the design of future exercise-based interventions for endometriosis pain management. These findings can inform exercise recommendations for endometriosis pain

management, specifically for targeting those who are at greater risk for sedentary behavior due to acute exacerbations in their pain after exercise.

Strengths and limitations of this study

- This study leveraged a mHealth-based design and participatory research to investigate daily exercise and pain symptom patterns in endometriosis under ecologically-valid conditions.
- The participant sample (N=1,009) represents 38 countries around the world, ages across the reproductive life span, and various socio-demographic conditions.
- The study is limited to self-report binary measure of exercise and did not have sufficient details on duration or intensity for inclusion in the analyses as potential moderators.
- Participants consisted of mostly white, non-Hispanic individuals and limited to somewhat consistent trackers, therefore results might not be generalizable to some demographic groups or less symptomatic endometriosis patients.

INTRODUCTION

Exercise, a subset of physical activity (PA) that is planned, structured, repetitive, and intended to improve or maintain physical fitness, is an important component of effective pain management.[1, 2] Both chronic (e.g., habitual) and acute (e.g., single session) exercise have been indicated to reduce pain and pain sensitivity (i.e., exercise-induced hypoalgesia).[3-5] Its efficacy for pain management has been demonstrated in numerous chronic pain conditions,[1, 6-10] with some reporting clinically meaningful reductions in pain severity associated with a range of exercise regimens.[8, 11] However, pain-related responses to exercise appear to be variable in populations with chronic pain conditions.[4] Similarly, exacerbation of pain with exercise could pose a barrier to regular exercise in such individuals, thus increasing resistance to exercising, which in return can worsen pain, related disability, and risk for co-morbidities.[12-14] Investigation into the naturally-occurring pattern of pain symptoms associated to exercise behavior can help inform the design of exercise-based therapies for targeting disease-related pain symptoms.

One population that can benefit from such investigation are individuals with endometriosis. Endometriosis is a systemic, estrogen-dependent inflammatory condition with debilitating symptoms including chronic pelvic pain, pain with sexual intercourse (dyspareunia), painful urination (dysuria), ovulation pain,[15-19] and is the second leading indication for hysterectomy.[20] There is substantial between-patient variation in its clinical manifestations,[17, 21] and a ~6.7(± 6.3) year-delay between symptom onset and its diagnosis[22]. It significantly impacts daily function and quality of life (QoL)[23, 24], contributing to a productivity loss of 6.3 hours/week[25] and an estimated $69.4 billion per year in excess health expenditures in the United States.[26] Despite its prevalence rate of 10% among

women of reproductive age,[17] endometriosis is still poorly understood[17, 21] and not well-studied,[27] with no cure. Existing medical and hormonal therapies have limited efficacy, often confounded by side effects.[28] Opioids and other analgesics are commonly prescribed to endometriosis patients[29, 30] despite lack of evidence for sufficient efficacy of their long-term use and serious side effects,[31, 32] as well as Centers for Disease Control and Prevention (CDC) guidelines recommending nonpharmacologic therapies, including PA.[31] These findings underscore the critical need to identify alternative approaches for endometriosis pain management.

One such approach is exercise, based on various mechanisms proposed in the literature[33] that might pertain to endometriosis. These include regulation of the serotonergic and opioid receptors,[34] reduction of inflammatory markers associated with pain,[35, 36] and exercise's effects on nerve growth factor expression that is associated with the painful endometriosis lesions.[37, 38] Exercise can increase pain management self-efficacy, a factor linked to improved pain outcomes and QoL in chronic pain.[39] While the evidence on exercise for pain management is promising [8, 40, 41], existing data are scarce, cross-sectional, and indicate variable effects.[41-45] There further is precedence to investigate whether habitual exercise frequency might moderate the association of acute exercise to pain symptoms. This is based on reported exercise-induced adaptations (i.e., habituation) to pain stimuli through increased pain threshold via involvement of the opioid system.[46, 47] Pain-related activation in the brain's descending antinociceptive pathway has been demonstrated among regular exercisers, with corresponding reductions in self-reported pain after acute bouts of at least moderate intensity exercise.[48] Moreover, habitual exercise levels have been indicated to moderate a variety of self-reported outcomes (e.g., mood, anxiety, fatigue) in response to acute exercise.[49-

51] While these findings are promising, their generalizability are limited by sample characteristics and laboratory-based experimental pain stimuli and exercise manipulations, and measurement duration. These collectively warrant further investigation into better understanding this relationship at a more granular level with a representative sample, under ecologically valid conditions, while accounting for possible between-individual variability.

Accordingly, this study investigates the association of exercise behavior to self-reported daily pain symptoms in endometriosis. We leverage mobile self-tracking, a particularly useful approach for capturing ecologically valid profiles of the dynamic temporal fluctuations and between-individual variability in pain over time.[52] We primarily aim to assess whether level of habitual exercise moderates the association of daily exercise to subsequent pain symptoms in endometriosis. Given the previously documented variable course of pain symptomology in endometriosis,[53] we also delineate the variability in day-to-day pain experiences within these analyses.

MATERIALS AND METHODS

Study design and protocols were approved by the Columbia University Irving Medical Center (CUIMC) Institutional Review Board (#AAAQ9812). This was an observational study conducted with retrospective data collected through a research-based smartphone self-tracking app designed and developed for tracking and documenting endometriosis and its self-management.

Study Setting: Phendo

Phendo is an observational research app available for iOS[1] and Android[2] for free in App stores. Phendo was designed using participatory design through a series of qualitative and

quantitative studies with endometriosis patients, [54, 55] with the goal of creating a patient-centered tool that engages the user as an active participant in the research on better understanding endometriosis.[56] As such, users of Phendo self-track as a form of participatory research, to contribute to creation of better documentation of the patient disease experience.[54, 57]

Informed Consent

Upon downloading Phendo and prior to starting to contribute their data, all participants are provided with an explanation of the App, purpose of the study (citizenendo.org), and provide formal electronic informed consent (and ascent for individuals 13-18 years old) (See Supplementary Figures 1-2 for example screenshots). Participants are instructed to track daily, but they are free to track as much or as sporadically as they wish, and they do not receive any prompts or requests to track a specific variable from the research team. Findings from a previous study evaluating recruitment and retention patterns within Phendo and across seven other research self-tracking apps for other diseases indicated that Phendo's engagement was on-par with standard engagement patterns in research smartphone apps.[58]

Study Sample

The study sample consisted of Phendo participants who reported an endometriosis diagnosis and had data on their daily exercise and pain. Endometriosis diagnosis in Phendo is determined based on participant response to the eligibility criteria item "Diagnosis: Do you have endometriosis?" with four possible options: "Yes, I was diagnosed as a result of surgery", "Yes, I was diagnosed by a medical professional without surgery", "I think I have endometriosis (know the symptoms/no doctor)", or "No". We a priori decided to include all participants who selected one of the three affirmative responses in the present analyses.

Patient and public involvement

Study measures in Phendo were previously developed through qualitative and quantitative studies (see [54, 57]) with participants with endometriosis, based on their habits, preferences, and needs for disease-specific symptom and activity tracking. This patient-centered participatory design technique is recommended for developing patient-reported outcome measures.[59-61] It has been suggested to enhance content validity[62], relevance to the target demographic and thus adherence to App use,[63, 64] therefore providing a more comprehensive and accurate representation of the relevant disease dimensions.[60]

Outcome Measures. Day-level pain was assessed through the following multiple-choice items within Phendo to capture all possible pain-related responses: 1. “Are you in pain now? Where is the pain?”, and 2. “Any gastrointestinal or urinary issues?”. For each item, severity was assessed using the item “How severe is the symptom?” with 3 options of mild, moderate, or severe. For the first item, the pain locations from which participants can select cover all areas of the body and organs (20 available choices, as well as right/left and upper/middle/lower specification), and can be mapped onto a visual analogous to the McGill Pain Scale.[65] The second item captures painful urination (dysuria), painful bowel movement (dyschezia). The severity question measures intensity on a 3-point categorical scale (mild, moderate, or severe), analogous to other commonly used pain rating scales in the literature.[66, 67] This discretization has been used for standardization and comparisons across different pain measures, demonstrated to better capture the nonlinear relationship between reported pain severity and interference with activity than by the use of numbers,[68] and circumvent the user-reported challenges of number-based intensity scales with respect to their range.[69]

While similar mHealth pain measurement approaches have been investigated for their validity, utility and specificity for various pain conditions[63, 69, 70], a standard "all-in-one" single outcome that captures the multi-dimensional pain experience across different populations remains to be established.[71, 72] To circumvent these issues, composite pain computations have been proposed.[73] We similarly computed a composite day-level pain score to capture participants' conceptualization of their pain experience based on area and severity.[53] It is computed heuristically by adding the severity scores reported for each body area (e.g., moderate pain in abdomen, mild pains in chest and leg would yield *2+1+1=4* as the total score). To account for and circumvent any potential pain rumination/catastrophizing [72, 74] and varying tracking habits among participants, the score was computed based on the unique reports of area-severity pairs per day for each participant (e.g., if a participant tracked mild abdominal pain three times in a day, this abdomen-mild pair is counted toward the daily pain score only once). This outcome measure was evaluated using two approaches in the analyses: 1) total pain score for the day, and 2) difference in this total score from previous day to the next (i.e., t-(t-1)). The second approach captures additional nuances in the data, enabling to distinguish between those with overall high pain scores over time and might experience a post-exercise reduction in pain versus those with low pain scores and does not experience a post-exercise reduction in pain.

Daily and habitual exercise. Exercise is tracked at the day level within the Phendo App with a binary (Yes/No) response to the root question "Did you exercise today?". Users can further customize their exercise tracking within their user profile, which are then saved for future tracking. This customizable item allows unrestricted free-text response, thus responses are highly variable. We relied on the root question to assess exercise at the day level and to compute weekly

exercise frequencies, and used the free-text entries to validate that the entries were exercise-related. This day-level assessment aims to increase ecological validity[59, 75] and reduce the likelihood of low test-retest reliability and inaccuracy due to recall bias.[76] Similar mHealth measures of daily PA and exercise have been used by others[77-79] who reported estimates in concordance with those from accelerometers,[80] showing higher correlations than do traditional self-report methods.[77, 78] Finally, our preliminary data (unpublished work) based on a sample of 30 Phendo users over the course of 14 days indicated significant associations (log odds ratio=1.44, z=3.00, p=0.002) of the Phendo exercise item responses to objectively-estimated and self-reported exercise levels.

Data Analysis

Sample Characteristics. We provide frequencies (%) and means (standard deviation; SD) for describing the study sample demographics. We characterize pain symptomology in the sample by describing the prevalence of pain severities by each body area.

Day-level associations of pain to exercise. We investigated the association of previous-day exercise to pain outcomes and the moderation of this association by habitual exercise levels using generalized linear mixed models (GLMMs). We estimated separate models predicting day-level total pain score and pain score difference. Both outcomes were regressed on previous-day exercise, habitual exercise levels, and their interaction to estimate the slope of average day-level pain and change in pain for each habitual exercise level. Participant as a random effect was included to account for between-person variability in daily pain by estimating a separate

intercept for each participant. Models were further adjusted for menstrual status(binary: yes/no), previous-day pain, body mass index (BMI) and education level.

Assessment of lagged-day effects (i.e., association of pain day *t* to exercise on day *t-1*) are motivated by 2 factors. First, this was necessary to ensure temporal sequence of the actual exercise and pain experienced by the participant. The App allows tracking of momentary pain through multiple daily entries, but allows tracking of exercise once a day. As such even if the participant exercises at multiple time points throughout the day, they are tracked together in a single daily entry. Second, though there is a plethora of literature on the acute exercise effects on a variety of health and disease outcomes (e.g., [81, 82]), studies are limited to measurements up to several hours. Investigation of an association between disease outcomes and previous-day exercise provides an opportunity to delineate possible sustained or lagged exercise effects.

Missing values in the variables were imputed as described in Supplementary File 1 (See Supplementary Table 1 and Supplementary Figures 3-5) and checked for appropriateness based on convergence and marginal distributions following guidelines.[83-85] We used a zero-inflated negative binomial (ZINB) distribution when modeling the total pain outcome, as it has been demonstrated to provide the best fit for outcomes with over-dispersion and zero-inflation,[86-88] as was the case for this variable. ZINB models consider two sources of zero observations: "sampling zeros" that are part of the underlying sampling distribution (i.e., negative binomial) and "structural zeros" that cannot score anything other than zero (i.e., participant did not track).[86] This virtue of the ZINB models allows for specification of the imputed zeros and prevents the risk of over-estimating effects and generates more conservative estimates for predictors of interest by estimating a separate zero-inflation term, as well as conditional model.[86] Given on our data inclusion approach (i.e., days with missing pain by default have a

tracked exercise or menstrual status response), we specified the zero-inflation term such that it was dependent on the exercise variable for the day, as well as assuming an overall general zero-inflation structure in the outcome through inclusion of an intercept, based on recommendations.[88] Menstrual status was not a significant predictor of zero inflation and therefore removed from the zero-inflation term during the modeling process. We included participants who had at least 11 pairs of consecutive days of data in the final analytic sample as this provided sufficient amount of data to 1) ensure model convergence and improve reliability and accuracy of the estimates, particularly the random effects and their variances[89-92], and 2) adequately infer participants' habitual exercise level by considering at least three weeks' worth of tracking to compute the weekly exercise frequency. All data analyses were conducted using R[93] and the glmmTMB package was used for the GLMMs.[87, 88]

RESULTS

Sample Descriptive Characteristics. Out of the initial eligible pool of 9,792 Phendo users with reported endometriosis, 7,949 had at least one day of tracking of the variables of interest for the study. Of these, 1,009 users had at least 11 pairs of consecutive days of data available on these variables and thus were included in the data analyses for the study. Sample characteristics are provided in Table 1. Participants had on average 89.6 days of data available for analysis (SD=62.8, Range=22-841, IQR=31). Tracked data span from November 2016 to April 2020. Participants collectively represent 38 countries around the world, with a wide age range (14-63 years), and varying education and employment status (See Table 1). Among participants, 702(69.5%) had laparoscopic confirmation of their diagnosis, 200(19.8%) had a clinician

diagnosis, and 107 (10.6%) had suspected endometriosis (i.e., "I think I have endometriosis (know the symptoms, no doctor)").

Description of pain symptomology. Mean daily pain score was 4.48 (SD=7.11, 0-79). Mean person-level daily pain score (i.e.,"mean of means") was 4.82 (SD=4.57, Range=0-34). Figure 1 depicts the prevalence of each pain severity per body area. Moderate intensity was the most frequently reported severity across all body areas (Mean=49.3%, SD=22.2), and pelvic pain was the most prevalent area, followed by back pain and gastrointestinal pain (See Figure 1).

Mean weekly exercise frequency was 1.43/week (SD=1.54, Range=0-6.87/week, IQR=2.21), 21.3% (N=215) of the sample had an exercise frequency of at least three times per week, and ~38.5% (388) of the sample did not engage in any regular exercise (i.e., <1/week). Consequently, ~40.2% (N=406) of the sample had an exercise frequency of 1-2 times per week. Prevalence of the 10 most frequently reported exercise modalities in the sample are depicted in Figure 2. Walking was the most common modality, reported by 50.94 % of the participants, followed by yoga (30.82%), and muscle strength/endurance training activities (24.38%). Yoga and stretching exercises were collectively reported by ~45% of the sample.

Association of day-level pain to exercise. Results of the GLMMs estimating day-level total pain score and difference are provided in Tables 2 and 3. Adequacy of imputations for valid statistical inference were verified based on the recommended measures of missing data information of *fraction of missing information* (λ) and *relative increase in variance due to nonresponse* (r)[94, 95] (See Supplemental File 2). Coefficients for the model interaction terms indicated a small but statistically significant moderation of previous-day exercise by habitual exercise levels (RR=0.96 for total pain score and -0.14 for pain score difference, $p<0.05$; See Figure 3). Participants with more frequent habitual exercise levels were more likely to report lower pain score and smaller

increases (or larger decreases) in pain the day after an exercise bout, compared to not having exercised the previous day. On the other hand, sedentary or less active individuals were more likely to report higher levels of pain and larger increases (or smaller decreases) in pain 1 day after an exercise bout compared to not having exercised the day before (See Table 1). Further inspection of this interaction indicated ~3 times/week of habitual exercise as the point after which previous day exercise began to be associated with favorable pain outcomes (e.g., a decrease from the predicted mean score) on the following day, adjusted for other day-level and person-level factors (Figure 3). There was substantial between-person variability in average day-level pain scores, based on the statistically significant random effect of participant in the models (See Tables 2 and 3, also depicted in Figure 4). The significance of this random effect can further be quantified through a restricted likelihood ratio test (RLRT) based on simulations from the model sample distribution, [96, 97] yielding an observed likelihood ratio (RLRT =7183.3, p-value < 0.0001). These collectively indicate substantial between-individual variability in daily pain experience contributing to the total model pain variance.

Post-hoc analyses. In a post-hoc analysis, we tested the possible influence of type of endometriosis diagnosis by including this categorical variable in the 2 models described above. Results indicated that diagnosis type did not have an influence on the results and were not significant predictors based on the non-significant B coefficients (p=0.48 and p=0.59 for pain score and p=0.70 and p=0.27 for difference in pain score) and that there were no differences across the 3 groups with respect to either daily total pain score or difference (χ^2 = 1415.1, df = 1438, p-value = 0.661) (See Supplementary Tables 2 and 3 for full results).

DISCUSSION

Leveraging mobile tracking to analyze 90,382 days of data from 1,009 women with endometriosis, this study investigated the association of exercise behavior to fluctuations in pain at the day level. In our analyses, the association of previous-day exercise to subsequent pain was moderated by habitual exercise levels. This effect was consistent across individuals independent of type of endometriosis diagnosis or BMI. There further was substantial between-person heterogeneity in naturally fluctuating pain patterns. To our knowledge, this is the first study to assess the association of day-level and habitual exercise to pain symptoms in endometriosis, and quantify the between-person heterogeneity in the natural fluctuations in pain in this population using a large sample of women around the world.

Moderation of the association of previous-day exercise to pain by habitual exercise levels suggest that, exercise behavior might first need to be developed as a sustained behavior (i.e., habit) to experience favorable pain outcomes associated with day-level exercise. Specifically, previous-day exercise was associated with more favorable pain outcomes when habitual exercise level reached 3/week in our sample. This is in line with the national PA guidelines [98], which recommend aerobic exercise at least 3/week and muscle-strengthening exercise at least 2/week.[99] However, there are no specific recommendations for endometriosis in the current guidelines; and systematic reviews recommend "overall, general exercise" without further details due to lack adequate research on the optimal dose of exercise for endometriosis pain.[8, 44] Our findings provide preliminary evidence for informing exercise recommendations for endometriosis pain management, specifically for targeting those who are at greater risk for sedentary behavior due to acute exacerbations in their pain after exercise.

Our findings on pain in pelvis and of moderate severity as the most frequently reported pain aspects are in line with those from others on endometriosis[100] and various chronic pain

conditions.[101, 102] The distribution of the total daily pain scores was right-skewed (i.e., extreme scores on the higher ends of the range) with a mean score that was on the lower end of the range. This could partly be due to the data collection method which includes not just days where the participant experienced pain but also days without pain. Indeed, our participants on average did not report or experience any pain 6.25% of the time. In contrast, traditional study designs typically rely on recall of past pain experience aggregated over a period of time (e.g., past week, month) and ask the participant to report their average or highest pain severity over this period.[103, 104] Such recall-based techniques are prone to peak-and-end effects,[105] and catastrophizing or other similar biases.[104, 106] Recruitment from clinical referral points is a common practice, and such patients are typically at the more disabled end of the spectrum. This has been attributed to higher normative scores in the literature,[103] as opposed to more even distributions of pain symptomology among community-based samples.[107] Self-tracking facilitates documentation of not only severe pain, but also mild, moderate, and no pain instances, therefore enabling a more realistic representation of the pain experience as it dynamically unfolds over time. This can reduce the likelihood of over-representing severe cases, which is a potential limitation attributed to data collected at point of contact in clinical settings.[23] However, it is difficult to make direct comparisons with other studies given the different pain measures, warranting further research.

The mean weekly exercise frequency was 1.43/week (SD=1.57, IQR=2.29) in the sample, with only 24.5 % (N=202) of the sample engaging in exercise at least three times a week. This suggests that individuals with endometriosis might be at increased risk for sedentary behavior or insufficient PA to meet the recommendations.[98, 99] Physical inactivity and sedentary behavior are risk factors for various comorbidities,[108] and have been linked to exacerbation of pain in

chronic pain populations.[109, 110] These collectively underscore the need to focus efforts on promoting regular physical activity in women with endometriosis. Though we did not analyze intensity, type or duration as potential model moderators, our findings suggest that there is a wide range of modalities preferred by this population, and that both aerobic and muscle strengthening and endurance type activities might be helpful. These modalities represent both lower and higher intensity ranges (e.g., yoga vs running/cycling), suggesting that responses to the exercise intensity might differ across individuals. Yoga and stretching were reported by almost half of the sample, which could indicate participants use these approaches for pain relief, in line with a previous study reporting efficacy of hatha yoga.[41] A Cochrane review concluded that exercise of ~50 minutes/session and at least three times per week may provide clinically significant reductions in menstrual pain regardless of the intensity.[8] However, authors noted the low quality of the existing evidence and a need for studies with larger, more diverse samples and appropriate control conditions.

Endometriosis patients are significantly more likely to have higher all-cause healthcare utilization and direct health care costs than controls, including twice the prevalence of opioid prescriptions for pain management (e.g., 77.2% vs 40% for endometriosis patients vs controls reported in one claims-based study),[30] and for prolonged durations (i.e., >90 days).[29] This is not recommended [33] as long-term use of opioids does not provide sufficient efficacy, and is associated with accidental overdose,[111] side effects such as gastrointestinal dysfunction,[31] and a paradoxical worsening of pain over time.[112] Exercise can further promote patient engagement, a recommended yet under-implemented component in chronic pain management[113] that can improve treatment adherence and outcomes. In line with our findings, substantial between-individual variability in exercise effects have been reported in the

literature.[41-44, 114] This can be targeted through individualized exercise prescriptions,[33, 115] providing evidence for undertaking a precision approach for self-management in endometriosis. Various individualization approaches have been investigated (e.g., adaptive treatment strategies,[116] micro-randomized trials,[117] just-in-time adaptive interventions [118]) for intervening on health behaviors and outcomes, including PA. [9, 117] It would be opportune to implement a similar N-of-1 intervention approach for identifying person-specific optimal "dose" of exercise based on its parameters (i.e., intensity, type, duration, frequency) to target endometriosis pain symptoms.

Another novel finding in our study was the lack of a difference in the pain experience, or in the association of previous day exercise and habitual exercise to subsequent pain outcomes between those with a formal- versus self-diagnosis of endometriosis. Endometriosis is difficult to diagnose, with a ~7.6 year delay with symptom onset and its surgical diagnosis.[26, 119, 120] Endometriosis patients further face insurance-related challenges in accessing healthcare for their condition.[16, 121] The participants without a formal diagnosis might have sought medical care for their symptoms but not received the needed care (e.g., diagnostic testing, referral to a specialist), or that their diagnostic tests results were false negative,[119] or alternatively did not have adequate access to healthcare. We refrain from making a conclusive remark, nevertheless; this finding underscores the need for further research in endometriosis conducted in diverse samples including possibly those self-report having endometriosis symptoms, instead of limiting to patients with a physician referral or simply relying on secondary data sources (e.g., electronic health records, claims databases).

We acknowledge several limitations of this study. First, we used a self-report binary measure of exercise in our analyses and did not have sufficient details on duration or intensity for

inclusion in the analyses as potential moderators. Similarly, the composite pain score has not been compared against existing standard pain measures in the literature for its validity, which is an area of investigation still under progress. Computation of a composite pain has been proposed by others[73] as this circumvents numerous limitations in current pain assessment approaches, including lack of a standard single outcome that can be used universally,[71] or a validated instrument that can capture all the constructs of persistent pain.[122] Similarly, there is a lack of endometriosis-specific pain measures for repeated assessment, and the categories of painful body locations/functions in this study are further reflective of how they are conceptualized and documented in traditional clinical records,[123] based on their mappings using standardized medical terminology nomenclature.[124] We relied on self-reported values for weight and height to compute BMI, which might have been under-estimated by some participants (e.g., those with higher BMI[125, 126]). Next, our sample consisted primarily of White, non-Hispanic women. Race/ethnicity was not significantly associated with average daily pain reports or exercise levels in the sample, based on the $\chi 2$ or Kruskal Wallis rank sum tests and therefore not included as a covariate in the models. Nevertheless, future studies are warranted to assess chronic pain in endometriosis measured over time across different racial/ethnic and socioeconomic groups. Similarly, these results are limited to users of the Phendo App and relatively consistent trackers, which might not be generalizable to those who do not actively track or monitor their diseases symptoms due to mildness of their disease and/or lack of interest in mHealth use.

Conclusion

In this study, we report habitual exercise levels as a potential moderator of the association of previous day exercise to endometriosis pain, suggesting that to accrue the benefits of exercise for adequate endometriosis pain management at the day level, exercise might first need to be

developed as a habit. While guidelines recommend prescribing exercise for management of pain in clinical populations, endometriosis (or general chronic) pain-specific recommendations to guide patients and providers on measurable parameters (time, type, intensity, and frequency) are lacking. This warrants future studies investigating the effects of both acute and chronic exercise on endometriosis pain with a focus on various types, intensities and durations.

Author Contributions

IE conceptualized the study, conducted the data analyses, and prepared the first draft of the manuscript. SLG and ENH were responsible for data acquisition, curation and management. NE acquired the funding and provided the mHealth infrastructure for the study (Phendo App). NE and SB provided guidance on the study design and data analyses. SB, NE, SLG and ENH reviewed and provided feedback on the the manuscript.

Funding

Funding for the work is provided by a postdoctoral fellowship from the Data Science Institute at Columbia University and an award from the National Library of Medicine (R01 LM013043). We are grateful to the Phendo participants.

Competing Interests

All authors report no conflicts of interest.

Data availability statement

Data are available on reasonable request.

1. Available at https://itunes.apple.com/us/app/phendo/id1145512423
2. Available at https://play.google.com/store/apps/details?id=com.appliedinformaticsinc.phendo

References

1. Ambrose KR, Golightly YM. Physical exercise as non-pharmacological treatment of chronic pain: Why and when. Best Pract Res Clin Rheumatol. 2015;29(1):120-30. Epub 2015/05/23. doi: 10.1016/j.berh.2015.04.022. PubMed PMID: 26267006.
2. Rice D, Nijs J, Kosek E, Wideman T, Hasenbring MI, Koltyn K, et al. Exercise-Induced Hypoalgesia in Pain-Free and Chronic Pain Populations: State of the Art and Future Directions. The Journal of Pain. 2019;20(11):1249-66. doi: https://doi.org/10.1016/j.jpain.2019.03.005.
3. Héroux M, Watt M, McGuire KA, Berardi JM. A personalized, multi-platform nutrition, exercise, and lifestyle coaching program: A pilot in women. Internet Interventions. 2017;7:16-22. doi: https://doi.org/10.1016/j.invent.2016.12.002.
4. Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH. Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. Cochrane Database of Systematic Reviews. 2017;(4).
5. Ambrose KR, Golightly YM. Physical exercise as non-pharmacological treatment of chronic pain: why and when. Best practice & research Clinical rheumatology. 2015;29(1):120-30.
6. Lockett D-MC, Campbell JF. The Effects of Aerobic Exercise on Migraine. Headache: The Journal of Head and Face Pain. 1992;32(1):50-4. doi: https://doi.org/10.1111/j.1526-4610.1992.hed3201050.x.
7. Lemmens J, De Pauw J, Van Soom T, Michiels S, Versijpt J, van Breda E, et al. The effect of aerobic exercise on the number of migraine days, duration and pain intensity in migraine: a systematic literature review and meta-analysis. J Headache Pain. 2019;20(1):16. Epub 2019/02/16. doi: 10.1186/s10194-019-0961-8. PubMed PMID: 30764753; PubMed Central PMCID: PMCPMC6734345.
8. Armour M, Ee CC, Naidoo D, Ayati Z, Chalmers KJ, Steel KA, et al. Exercise for dysmenorrhoea. Cochrane Database of Systematic Reviews. 2019;(9). doi: 10.1002/14651858.CD004142.pub4. PubMed PMID: CD004142.
9. Rabbi M, Aung MS, Gay G, Reid MC, Choudhury T. Feasibility and acceptability of mobile phone–based auto-personalized physical activity recommendations for chronic pain self-management: pilot study on adults. Journal of medical Internet research. 2018;20(10):e10147.
10. Sevel L, Boissoneault J, Alappattu M, Bishop M, Robinson M. Training endogenous pain modulation: a preliminary investigation of neural adaptation following repeated exposure to clinically-relevant pain. Brain Imaging and Behavior. 2020;14(3):881-96. doi: 10.1007/s11682-018-0033-8.
11. Gordon R, Bloxham S, editors. A systematic review of the effects of exercise and physical activity on non-specific chronic low back pain. Healthcare; 2016: Multidisciplinary Digital Publishing Institute.
12. Zhang R, Chomistek AK, Dimitrakoff JD, Giovannucci EL, Willett WC, Rosner BA, et al. Physical activity and chronic prostatitis/chronic pelvic pain syndrome. Med Sci Sports Exerc. 2015;47(4):757-64. doi: 10.1249/MSS.0000000000000472. PubMed PMID: 25116086.

13. Pinto A, Di Raimondo D, Tuttolomondo A, Buttà C, Milio G, Licata G. Effects of physical exercise on inflammatory markers of atherosclerosis. Current pharmaceutical design. 2012;18(28):4326-49.
14. Garatachea N, Molinero O, Martínez-García R, Jimenez-Jimenez R, Gonzalez-Gallego J, Marquez S. Feelings of well being in elderly people: relationship to physical activity and physical function. Archives of Gerontology and Geriatrics. 2009;48(3):306-12.
15. Tamaresis JS, Irwin JC, Goldfien GA, Rabban JT, Burney RO, Nezhat C, et al. Molecular classification of endometriosis and disease stage using high-dimensional genomic data. Endocrinology. 2014;155(12):4986-99. Epub 2014/09/23. doi: 10.1210/en.2014-1490. PubMed PMID: 25243856; PubMed Central PMCID: PMCPMC4239429.
16. Fourquet J, Gao X, Zavala D, Orengo JC, Abac S, Ruiz A, et al. Patients' report on how endometriosis affects health, work, and daily life. Fertil Steril. 2010;93(7):2424-8. doi: 10.1016/j.fertnstert.2009.09.017. PubMed PMID: 19926084.
17. Rogers PA, D'Hooghe TM, Fazleabas A, Gargett CE, Giudice LC, Montgomery GW, et al. Priorities for endometriosis research: recommendations from an international consensus workshop. Reprod Sci. 2009;16(4):335-46. Epub 2009/02/07. doi: 10.1177/1933719108330568. PubMed PMID: 19196878; PubMed Central PMCID: PMCPMC3682634.
18. Schliep KC, Mumford SL, Peterson CM, Chen Z, Johnstone EB, Sharp HT, et al. Pain typology and incident endometriosis. Hum Reprod. 2015;30(10):2427-38. Epub 2015/08/11. doi: 10.1093/humrep/dev147. PubMed PMID: 26269529.
19. Fourquet J, Zavala DE, Missmer S, Bracero N, Romaguera J, Flores I. Disparities in healthcare services in women with endometriosis with public vs. private health insurance. Am J Obstet Gynecol. 2019. Epub 2019/06/22. doi: 10.1016/j.ajog.2019.06.020. PubMed PMID: 31226295.
20. Whiteman MK, Hillis SD, Jamieson DJ, Morrow B, Podgornik MN, Brett KM, et al. Inpatient hysterectomy surveillance in the United States, 2000-2004. American journal of obstetrics and gynecology. 2008;198(1):34. e1-. e7.
21. Garry R. The endometriosis syndromes: a clinical classification in the presence of aetiological confusion and therapeutic anarchy. Human Reproduction. 2004;19(4):760-8. doi: 10.1093/humrep/deh147.
22. Nnoaham KE, Hummelshoj L, Webster P, d'Hooghe T, de Cicco Nardone F, de Cicco Nardone C, et al. Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries. Fertil Steril. 2011;96(2):366-73 e8. Epub 2011/07/02. doi: 10.1016/j.fertnstert.2011.05.090. PubMed PMID: 21718982; PubMed Central PMCID: PMCPMC3679489.
23. De Graaff A, D'hooghe T, Dunselman G, Dirksen C, Hummelshoj L, Consortium WE, et al. The significant effect of endometriosis on physical, mental and social wellbeing: results from an international cross-sectional survey. Human reproduction. 2013;28(10):2677-85.
24. Simoens S, Dunselman G, Dirksen C, Hummelshoj L, Bokor A, Brandes I, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. Human Reproduction. 2012;27(5):1292-9.
25. Soliman AM, Coyne KS, Gries KS, Castelli-Haley J, Snabes MC, Surrey ES. The effect of endometriosis symptoms on absenteeism and presenteeism in the workplace and at home. Journal of managed care & specialty pharmacy. 2017;23(7):745-54.
26. Simoens S, Dunselman G, Dirksen C, Hummelshoj L, Bokor A, Brandes I, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in

referral centres. Hum Reprod. 2012;27(5):1292-9. Epub 2012/03/17. doi: 10.1093/humrep/des073. PubMed PMID: 22422778.
27. Health. NIo. Estimates of funding for various research, condition, and disease categories (RCDC). 2019 [cited 2019 October 9]. Available from: https://report.nih.gov/categorical_spending.aspx.
28. The Practice Committee of the American Society for Reproductive Medicine. Treatment of pelvic pain associated with endometriosis: a committee opinion. 2014 2014/04/01/. Report No.: 0015-0282 Contract No.: 4.
29. Lamvu G, Soliman AM, Manthena SR, Gordon K, Knight J, Taylor HS. Patterns of prescription opioid use in women with endometriosis: evaluating prolonged use, daily dose, and concomitant use with benzodiazepines. Obstetrics and gynecology. 2019;133(6):1120.
30. Soliman AM, Surrey ES, Bonafede M, Nelson JK, Vora JB, Agarwal SK. Health care utilization and costs associated with endometriosis among women with medicaid insurance. Journal of managed care & specialty pharmacy. 2019;25(5):566-72.
31. Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain—United States, 2016. Jama. 2016;315(15):1624-45.
32. Krebs EE, Gravely A, Nugent S, Jensen AC, DeRonne B, Goldsmith ES, et al. Effect of opioid vs nonopioid medications on pain-related function in patients with chronic back pain or hip or knee osteoarthritis pain: the SPACE randomized clinical trial. Jama. 2018;319(9):872-82.
33. Sluka KA, Frey-Law L, Hoeger Bement M. Exercise-induced pain and analgesia? Underlying mechanisms and clinical translation. Pain. 2018;159 Suppl 1(Suppl 1):S91-S7. doi: 10.1097/j.pain.0000000000001235. PubMed PMID: 30113953.
34. Tour J, Löfgren M, Mannerkorpi K, Gerdle B, Larsson A, Palstam A, et al. Gene-to-gene interactions regulate endogenous pain modulation in fibromyalgia patients and healthy controls-antagonistic effects between opioid and serotonin-related genes. Pain. 2017;158(7):1194-203. Epub 2017/03/11. doi: 10.1097/j.pain.0000000000000896. PubMed PMID: 28282362; PubMed Central PMCID: PMCPMC5472004.
35. Bobinski F, Teixeira JM, Sluka KA, Santos ARS. Interleukin-4 mediates the analgesia produced by low-intensity exercise in mice with neuropathic pain. Pain. 2018;159(3):437-50. Epub 2017/11/16. doi: 10.1097/j.pain.0000000000001109. PubMed PMID: 29140923; PubMed Central PMCID: PMCPMC5812806.
36. Montenegro ML, Bonocher CM, Meola J, Portella RL, Ribeiro-Silva A, Brunaldi MO, et al. Effect of Physical Exercise on Endometriosis Experimentally Induced in Rats. Reproductive sciences (Thousand Oaks, Calif). 2018:1933719118799205. Epub 2018/09/21. doi: 10.1177/1933719118799205. PubMed PMID: 30231769.
37. Stratton P, Berkley KJ. Chronic pelvic pain and endometriosis: translational evidence of the relationship and implications. Human reproduction update. 2011;17(3):327-46. Epub 2010/11/23. doi: 10.1093/humupd/dmq050. PubMed PMID: 21106492.
38. Park S-J, Yong M-S, Na S-S. Effect of exercise on the expression of nerve growth factor in the spinal cord of rats with induced osteoarthritis. Journal of physical therapy science. 2015;27(8):2551-4. Epub 2015/08/21. doi: 10.1589/jpts.27.2551. PubMed PMID: 26357438.
39. Karasawa Y, Yamada K, Iseki M, Yamaguchi M, Murakami Y, Tamagawa T, et al. Association between change in self-efficacy and reduction in disability among patients with chronic pain. PLOS ONE. 2019;14(4):e0215404. doi: 10.1371/journal.pone.0215404.

40. Armour M, Sinclair J, Chalmers KJ, Smith CA. Self-management strategies amongst Australian women with endometriosis: a national online survey. BMC Complementary and Alternative Medicine. 2019;19(1):17. doi: 10.1186/s12906-019-2431-x.
41. Gonçalves AV, Barros NF, Bahamondes L. The Practice of Hatha Yoga for the Treatment of Pain Associated with Endometriosis. Journal of Alternative & Complementary Medicine. 2017;23(1):45-52. doi: 10.1089/acm.2015.0343. PubMed PMID: 120746246.
42. Ricci E, Viganò P, Cipriani S, Chiaffarino F, Bianchi S, Rebonato G, et al. Physical activity and endometriosis risk in women with infertility or pain: Systematic review and meta-analysis. Medicine. 2016;95(40):e4957-e. doi: 10.1097/MD.0000000000004957. PubMed PMID: 27749551.
43. Carpenter SE, Tjaden B, Rock JA, Kimball A. The effect of regular exercise on women receiving danazol for treatment of endometriosis. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics. 1995;49(3):299-304. Epub 1995/06/01. PubMed PMID: 9764869.
44. Bonocher CM, Montenegro ML, Rosa ESJC, Ferriani RA, Meola J. Endometriosis and physical exercises: a systematic review. Reproductive biology and endocrinology : RB&E. 2014;12:4. Epub 2014/01/08. doi: 10.1186/1477-7827-12-4. PubMed PMID: 24393293; PubMed Central PMCID: PMCPMC3895811.
45. Naugle KM, Fillingim RB, Riley JL, 3rd. A meta-analytic review of the hypoalgesic effects of exercise. The journal of pain : official journal of the American Pain Society. 2012;13(12):1139-50. Epub 2012/11/08. doi: 10.1016/j.jpain.2012.09.006. PubMed PMID: 23141188.
46. Janal MN, Colt EWD, Clark WC, Glusman M. Pain sensitivity, mood and plasma endocrine levels in man following long-distance running: Effects of naloxone. Pain. 1984;19(1):13-25. doi: https://doi.org/10.1016/0304-3959(84)90061-7.
47. Droste C, Greenlee MW, Schreck M, Roskamm H. Experimental pain thresholds and plasma beta-endorphin levels during exercise. Medicine & Science in Sports & Exercise. 1991;23(3):334-42. doi: 10.1249/00005768-199103000-00012.
48. Scheef L, Jankowski J, Daamen M, Weyer G, Klingenberg M, Renner J, et al. An fMRI study on the acute effects of exercise on pain processing in trained athletes. PAIN. 2012;153(8).
49. Hoffman MD, Hoffman DR. Exercisers Achieve Greater Acute Exercise-Induced Mood Enhancement Than Nonexercisers. Archives of Physical Medicine and Rehabilitation. 2008;89(2):358-63. doi: 10.1016/j.apmr.2007.09.026.
50. Hallgren M, Moss ND, Gastin P. Regular exercise participation mediates the affective response to acute bouts of vigorous exercise. J Sports Sci Med. 2010;9(4):629-37. Epub 2010/01/01. PubMed PMID: 24149790; PubMed Central PMCID: PMCPMC3761821.
51. Chen Y-C, Chen C, Martínez RM, Etnier JL, Cheng Y. Habitual physical activity mediates the acute exercise-induced modulation of anxiety-related amygdala functional connectivity. Scientific Reports. 2019;9(1):19787. doi: 10.1038/s41598-019-56226-z.
52. May M, Junghaenel DU, Ono M, Stone AA, Schneider S. Ecological Momentary Assessment Methodology in Chronic Pain Research: A Systematic Review. J Pain. 2018;19(7):699-716. Epub 2018/01/31. doi: 10.1016/j.jpain.2018.01.006. PubMed PMID: 29371113.
53. Ensari I, Pichon A, Lipsky-Gorman S, Bakken S, Elhadad N. Augmenting the Clinical Data Sources for Enigmatic Diseases: A Cross-Sectional Study of Self-Tracking Data and Clinical Documentation in Endometriosis. Applied Clinical Informatics. 2020;11(05):769-84.

54. McKillop M, Mamykina L, Elhadad N, editors. Designing in the Dark: Eliciting Self-tracking Dimensions for Understanding Enigmatic Disease. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; 2018: ACM.
55. Urteaga I, McKillop M, Elhadad N. Learning endometriosis phenotypes from patient-generated data. npj Digital Medicine. 2020;3(1):88. doi: 10.1038/s41746-020-0292-9.
56. Ensari I, Elhadad N. Chapter 5 - mHealth for research: participatory research applications to gain disease insights. In: Syed-Abdul S, Zhu X, Fernandez-Luque L, editors. Digital Health: Elsevier; 2021. p. 79-102.
57. McKillop M, Voigt N, Schnall R, Elhadad N. Exploring self-tracking as a participatory research activity among women with endometriosis. Journal of Participatory Medicine. 2016.
58. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. npj Digital Medicine. 2020;3(1):21. doi: 10.1038/s41746-020-0224-8.
59. Frost MH, Reeve BB, Liepa AM, Stauffer JW, Hays RD. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? Value Health. 2007;10 Suppl 2:S94-s105. Epub 2007/11/13. doi: 10.1111/j.1524-4733.2007.00272.x. PubMed PMID: 17995479.
60. Anthoine E, Moret L, Regnault A, Sébille V, Hardouin J-B. Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. Health Qual Life Outcomes. 2014;12:176-. doi: 10.1186/s12955-014-0176-2. PubMed PMID: 25492701.
61. US Department of Health Human Services. Guidance for industry-Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.
62. Lomas J, Pickard L, Mohide A. Patient versus clinician item generation for quality-of-life measures: the case of language-disabled adults. Medical Care. 1987:764-9.
63. Jamison RN, Raymond SA, Levine JG, Slawsby EA, Nedeljkovic SS, Katz NP. Electronic diaries for monitoring chronic pain: 1-year validation study. Pain. 2001;91(3):277-85.
64. Haythornthwaite JA, Menefee LA, Heinberg LJ, Clark MR. Pain coping strategies predict perceived control over pain. Pain. 1998;77(1):33-9.
65. Melzack R. The McGill Pain Questionnaire: Major properties and scoring methods. PAIN. 1975;1(3).
66. Jones KR, Vojir CP, Hutt E, Fink R. Determining mild, moderate, and severe pain equivalency across pain-intensity tools in nursing home residents. J Rehabil Res Dev. 2007;44(2):305-14. Epub 2007/06/07. doi: 10.1682/jrrd.2006.05.0051. PubMed PMID: 17551881.
67. Bestel E, Gotteland J-P, Donnez J, Taylor RN, Garner EI. Linzagolix for Endometriosis-Associated Pain: Lipid Changes After 52 Weeks of Treatment [25B]. Obstetrics & Gynecology. 2020;135:25S. doi: 10.1097/01.AOG.0000663180.46470.c9. PubMed PMID: 00006250-202005001-00082.
68. Serlin RC, Mendoza TR, Nakamura Y, Edwards KR, Cleeland CS. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. Pain. 1995;61(2):277-84. Epub 1995/05/01. doi: 10.1016/0304-3959(94)00178-h. PubMed PMID: 7659438.
69. Adams P, Murnane EL, Elfenbein M, Wethington E, Gay G. Supporting the Self-Management of Chronic Pain Conditions with Tailored Momentary Self-Assessments. Proc

SIGCHI Conf Hum Factor Comput Syst. 2017;2017:1065-77. doi: 10.1145/3025453.3025832. PubMed PMID: 30310887.
70. Lee RR, Rashid A, Ghio D, Thomson W, Cordingley L. "Seeing Pain Differently": A Qualitative Investigation Into the Differences and Similarities of Pain and Rheumatology Specialists' Interpretation of Multidimensional Mobile Health Pain Data From Children and Young People With Juvenile Idiopathic Arthritis. JMIR Mhealth Uhealth. 2019;7(7):e12952. Epub 2019/07/04. doi: 10.2196/12952. PubMed PMID: 31267979; PubMed Central PMCID: PMCPMC6632104.
71. Bouhassira D, Attal N. All in one: Is it possible to assess all dimensions of any pain with a simple questionnaire? PAIN. 2009;144(1).
72. Boonstra AM, Stewart RE, Köke AJA, Oosterwijk RFA, Swaan JL, Schreurs KMG, et al. Cut-Off Points for Mild, Moderate, and Severe Pain on the Numeric Rating Scale for Pain in Patients with Chronic Musculoskeletal Pain: Variability and Influence of Sex and Catastrophizing. Front Psychol. 2016;7:1466-. doi: 10.3389/fpsyg.2016.01466. PubMed PMID: 27746750.
73. Pilitsis JG, Fahey M, Custozzo A, Chakravarthy K, Capobianco R. Composite score is a better reflection of patient response to chronic pain therapy compared with pain intensity alone. Neuromodulation: Technology at the Neural Interface. 2021;24(1):68-75.
74. Dirks JF, Wunder J, Kinsman R, McElhinny J, Jones NF. A Pain Rating Scale and a Pain Behavior Checklist for Clinical Use: Development, Norms, and the Consistency Score. Psychotherapy and Psychosomatics. 1993;59(1):41-9.
75. Faurholt-Jepsen M, Munkholm K, Frost M, Bardram JE, Kessing LV. Electronic self-monitoring of mood using IT platforms in adult patients with bipolar disorder: A systematic review of the validity and evidence. BMC Psychiatry. 2016;16(1):7. doi: 10.1186/s12888-016-0713-0.
76. Charter RA. Sample size requirements for precise estimates of reliability, generalizability, and validity coefficients. J Clin Exp Neuropsychol. 1999;21(4):559-66. Epub 1999/11/07. doi: 10.1076/jcen.21.4.559.889. PubMed PMID: 10550813.
77. Knell G, Gabriel KP, Businelle MS, Shuval K, Wetter DW, Kendzor DE. Ecological Momentary Assessment of Physical Activity: Validation Study. J Med Internet Res. 2017;19(7):e253. doi: 10.2196/jmir.7602.
78. Swendeman D, Comulada WS, Koussa M, Worthman CM, Estrin D, Rotheram-Borus MJ, et al. Longitudinal Validity and Reliability of Brief Smartphone Self-Monitoring of Diet, Stress, and Physical Activity in a Diverse Sample of Mothers. JMIR Mhealth Uhealth. 2018;6(9):e176. Epub 2018/09/27. doi: 10.2196/mhealth.9378. PubMed PMID: 30249576; PubMed Central PMCID: PMCPMC6231816.
79. Katapally TR, Chu LM. Digital epidemiological and citizen science methodology to capture prospective physical activity in free-living conditions: a SMART Platform study. BMJ Open. 2020;10(6):e036787. Epub 2020/07/01. doi: 10.1136/bmjopen-2020-036787. PubMed PMID: 32595163; PubMed Central PMCID: PMCPMC7322321.
80. Zink J, Belcher BR, Dzubur E, Ke W, O'Connor S, Huh J, et al. Association Between Self-Reported and Objective Activity Levels by Demographic Factors: Ecological Momentary Assessment Study in Children. JMIR Mhealth Uhealth. 2018;6(6):e150. doi: 10.2196/mhealth.9592.

81. Ensari I, Greenlee TA, Motl RW, Petruzzello SJ. Meta-analysis of acute exercise effects on state anxiety: An update of randomized controlled trials over the past 25 years. Depression and anxiety. 2015;32(8):624-34.
82. Ensari I, Petruzzello SJ, Motl RW. The effects of acute yoga on anxiety symptoms in response to a carbon dioxide inhalation task in women. Complementary Therapies in Medicine. 2019:102230. doi: https://doi.org/10.1016/j.ctim.2019.102230.
83. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software; Vol 1, Issue 3 (2011). 2011.
84. Van Buuren S. Flexible imputation of missing data: CRC press; 2018.
85. Bondarenko I, Raghunathan T. Graphical and numerical diagnostic tools to assess suitability of multiple imputations and imputation models. Statistics in Medicine. 2016;35(17):3007-20. doi: https://doi.org/10.1002/sim.6926.
86. Hu M-C, Pavlicova M, Nunes EV. Zero-inflated and hurdle models of count data with extra zeros: examples from an HIV-risk reduction intervention trial. Am J Drug Alcohol Abuse. 2011;37(5):367-75. doi: 10.3109/00952990.2011.597280. PubMed PMID: 21854279.
87. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R journal. 2017;9(2):378-400.
88. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. Modeling zero-inflated count data with glmmTMB. bioRxiv. 2017:132753. doi: 10.1101/132753.
89. Schunck R. Cluster Size and Aggregated Level 2 Variables in Multilevel Models. A Cautionary Note. 2016. 2016;10(1). Epub 2016-07-20. doi: 10.12758/mda.2016.005.
90. Bell B, Ferron J, Kromrey J, editors. Cluster Size in Multilevel Models: The Impact of Sparse Data Structures on Point and Interval Estimates in Two-Level Models2008.
91. Austin PC, Leckie G. The effect of number of clusters and cluster size on statistical power and Type I error rates when testing random effects variance components in multilevel linear and logistic regression models. Journal of Statistical Computation and Simulation. 2018;88(16):3151-63. doi: 10.1080/00949655.2018.1504945.
92. Everitt B, Howell D, Snijders T, editors. Power and sample size in multilevel modeling2006.
93. Team R. Core (2020). R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria URL https://www R-project org. 2020.
94. Rubin DB. The Calculation of Posterior Distributions by Data Augmentation: Comment: A Noniterative Sampling/Importance Resampling Alternative to the Data Augmentation Algorithm for Creating a Few Imputations When Fractions of Missing Information Are Modest: The SIR Algorithm. Journal of the American Statistical Association. 1987;82(398):543-6. doi: 10.2307/2289460.
95. Rubin DB. Multiple imputation for nonresponse in surveys: John Wiley & Sons; 2004.
96. Scheipl F, Greven S, Küchenhoff H. Size and power of tests for a zero random effect variance or polynomial regression in additive and linear mixed models. Computational statistics & data analysis. 2008;52(7):3283-99.
97. Crainiceanu CM, Ruppert D. Likelihood ratio tests in linear mixed models with one variance component. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2004;66(1):165-85.
98. US Department of Health Human Services Physical activity guidelines advisory committee scientific report. 2018.

99. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. Jama. 2018;320(19):2020-8.
100. Warzecha D, Szymusik I, Wielgos M, Pietrzak B. The Impact of Endometriosis on the Quality of Life and the Incidence of Depression-A Cohort Study. Int J Environ Res Public Health. 2020;17(10). Epub 2020/05/28. doi: 10.3390/ijerph17103641. PubMed PMID: 32455821; PubMed Central PMCID: PMCPMC7277332.
101. Becker N, Thomsen AB, Olsen AK, Sjogren P, Bech P, Eriksen J. Pain epidemiology and health related quality of life in chronic non-malignant pain patients referred to a Danish multidisciplinary pain center. Pain. 1997;73(3):393-400. doi: 10.1016/s0304-3959(97)00126-7. PubMed PMID: WOS:000071429200014.
102. Bouhassira D, Lantéri-Minet M, Attal N, Laurent B, Touboul C. Prevalence of chronic pain with neuropathic characteristics in the general population. Pain. 2008;136(3):380-7.
103. Nicholas MK, Asghari A, Blyth FM. What do the numbers mean? Normative data in chronic pain measures. Pain. 2008;134(1):158-73. doi: https://doi.org/10.1016/j.pain.2007.04.007.
104. Dansie EJ, Turk DC. Assessment of patients with chronic pain. Br J Anaesth. 2013;111(1):19-25. doi: 10.1093/bja/aet124. PubMed PMID: 23794641.
105. Schneider S, Stone AA, Schwartz JE, Broderick JE. Peak and end effects in patients' daily recall of pain and fatigue: a within-subjects analysis. J Pain. 2011;12(2):228-35. doi: 10.1016/j.jpain.2010.07.001. PubMed PMID: 20817615.
106. De Boer M, Struys M, Versteegen G. Pain-related catastrophizing in pain patients and people with pain in the general population. European journal of pain. 2012;16(7):1044-52.
107. Ehde DM, Gibbons LE, Chwastiak L, Bombardier CH, Sullivan MD, Kraft GH. Chronic pain in a large community sample of persons with multiple sclerosis. Multiple Sclerosis Journal. 2003;9(6):605-11. doi: 10.1191/1352458503ms939oa.
108. Katzmarzyk PT, Powell KE, Jakicic JM, Troiano RP, Piercy K, Tennant B, et al. Sedentary Behavior and Health: Update from the 2018 Physical Activity Guidelines Advisory Committee. Med Sci Sports Exerc. 2019;51(6):1227-41. doi: 10.1249/MSS.0000000000001935. PubMed PMID: 31095080.
109. Basen-Engquist K, Scruggs S, Jhingran A, Bodurka DC, Lu K, Ramondetta L, et al. Physical activity and obesity in endometrial cancer survivors: associations with pain, fatigue, and physical functioning. Am J Obstet Gynecol. 2009;200(3):288.e1-.e2888. Epub 2008/12/25. doi: 10.1016/j.ajog.2008.10.010. PubMed PMID: 19110220.
110. Dansie EJ, Turk DC, Martin KR, Van Domelen DR, Patel KV. Association of Chronic Widespread Pain With Objectively Measured Physical Activity in Adults: Findings From the National Health and Nutrition Examination Survey. The Journal of Pain. 2014;15(5):507-15. doi: https://doi.org/10.1016/j.jpain.2014.01.489.
111. Just JM, Scherbaum N, Specka M, Puth MT, Weckbecker K. Rate of opioid use disorder in adults who received prescription opioid pain therapy-A secondary data analysis. PLoS One. 2020;15(7):e0236268. Epub 2020/07/24. doi: 10.1371/journal.pone.0236268. PubMed PMID: 32702036.
112. Sinatra R. Causes and consequences of inadequate management of acute pain. Pain Med. 2010;11(12):1859-71. Epub 2010/11/03. doi: 10.1111/j.1526-4637.2010.00983.x. PubMed PMID: 21040438.

113. Engeler DS, Baranowski AP, Dinis-Oliveira P, Elneil S, Hughes J, Messelink EJ, et al. The 2013 EAU guidelines on chronic pelvic pain: is management of chronic pelvic pain a habit, a philosophy, or a science? 10 years of development. European urology. 2013;64(3):431-9.
114. Awad E, Ahmed HAH, Yousef A, Abbas R. Efficacy of exercise on pelvic pain and posture associated with endometriosis: within subject design. Journal of physical therapy science. 2017;29(12):2112-5. Epub 2017/12/07. doi: 10.1589/jpts.29.2112. PubMed PMID: 29643586.
115. Polaski AM, Phelps AL, Kostek MC, Szucs KA, Kolber BJ. Exercise-induced hypoalgesia: A meta-analysis of exercise dosing for the treatment of chronic pain. PloS one. 2019;14(1):e0210418-e. doi: 10.1371/journal.pone.0210418. PubMed PMID: 30625201.
116. Almirall D, Compton SN, Gunlicks-Stoessel M, Duan N, Murphy SA. Designing a pilot sequential multiple assignment randomized trial for developing an adaptive treatment strategy. Statistics in medicine. 2012;31(17):1887-902. Epub 2012/03/23. doi: 10.1002/sim.4512. PubMed PMID: 22438190; PubMed Central PMCID: PMCPMC3399974.
117. Klasnja P, Smith S, Seewald NJ, Lee A, Hall K, Luers B, et al. Efficacy of Contextually Tailored Suggestions for Physical Activity: A Micro-randomized Optimization Trial of HeartSteps. Annals of behavioral medicine : a publication of the Society of Behavioral Medicine. 2018. Epub 2018/09/08. doi: 10.1093/abm/kay067. PubMed PMID: 30192907.
118. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. Annals of behavioral medicine : a publication of the Society of Behavioral Medicine. 2018;52(6):446-62. Epub 2016/09/25. doi: 10.1007/s12160-016-9830-8. PubMed PMID: 27663578; PubMed Central PMCID: PMCPMC5364076.
119. Falcone T, Mascha E. The elusive diagnostic test for endometriosis. Fertility and sterility. 2003;80(4):886-8.
120. Marian S, Hermanowicz-Szamatowicz K. Endometriosis–a decade later–still an enigmatic disease. What is the new in the diagnosis and treatment? Gynecological Endocrinology. 2020;36(2):104-8.
121. Fourquet J, Zavala DE, Missmer S, Bracero N, Romaguera J, Flores I. Disparities in healthcare services in women with endometriosis with public vs private health insurance. American Journal of Obstetrics & Gynecology. 2019;221(6):623.e1-.e11. doi: 10.1016/j.ajog.2019.06.020.
122. Grimmer-Somers K, Vipond N, Kumar S, Hall G. A review and critique of assessment instruments for patients with persistent pain. Journal of pain research. 2009;2:21.
123. Ensari I, Pichon A, Lipsky-Gorman S, Bakken S, Elhadad N. Augmenting the Clinical Data Sources for Enigmatic Diseases: a Cross-Sectional Study of Self-Tracking Data and Clinical Documentation in Endometriosis Applied Clinical Informatics, in press. 2020.
124. Donnelly K. SNOMED-CT: The advanced terminology and coding system for eHealth. Studies in health technology and informatics. 2006;121:279.
125. Gillum RF, Sempos CT. Ethnic variation in validity of classification of overweight and obesity using self-reported weight and height in American women and men: the Third National Health and Nutrition Examination Survey. Nutrition journal. 2005;4(1):27.
126. Larsen JK, Ouwens M, Engels RC, Eisinga R, van Strien T. Validity of self-reported weight and height and predictors of weight bias in female college students. Appetite. 2008;50(2-3):386-9.

127. Grund S, Lüdtke O, Robitzsch A. Multiple imputation of missing data for multilevel models: Simulations and recommendations. Organizational Research Methods. 2018;21(1):111-49.

128. Barnard J, Rubin DB. Small-Sample Degrees of Freedom with Multiple Imputation. Biometrika. 1999;86(4):948-55.

129. Little RJ. Missing-data adjustments in large surveys. Journal of Business & Economic Statistics. 1988;6(3):287-96.

130. Nguyen CD, Carlin JB, Lee KJ. Model checking in multiple imputation: an overview and case study. Emerging Themes in Epidemiology. 2017;14(1):8. doi: 10.1186/s12982-017-0062-6.

Table 1. Study Sample Characteristics (N=1009).

Characteristic (N)	Mean (SD)
Age (827)	31.0 (7.26), Median=30.6 (MAD=7.41), Range= 14.3-62.9
BMI (787)	25.9 (6.98), Median=24.1 (MAD=4.74), Range= 16.01-72.24
Type of endometriosis diagnosis	Frequency (%)
Surgery	702 (69.57%)
Clinician	200 (19.82%)
Self-diagnosis	107 (10.60%)
Characteristic (N)	**Frequency (%)**
Work Environment	
Home (218)	26.42
Outside (570)	69.09
Unknown (221)	21.29
Living environment	
Rural (129)	15.27
Suburban (340)	41.21
Urban (363)	44.00
Unknown (161)	19.5
Relationship status	
Married/domestic partnership (442)	53.57
Separated/divorced (28)	3.39
Single/never married (310)	37.57

Unknown (229)	22.69
Education Level	
College or higher (547)	66.30
High school graduate or less (74)	8.96
Some college (209)	25.33
Unknown (179)	17.7
Employment Status	
Employed (541)	65.57
Not employed (120)	14.54
Student (129)	15.63
Unknown (219)	21.70
Race/Ethnicity	
White, Non-Hispanic (699)	84.72
Black, Non-Hispanic (20)	2.42
Asian (22)	2.6
Native American (6)	0.72
Hispanic (38)	4.6
Other (51)	6.18
Unknown (173)	17.14
Country of Residence	
United States (444)	44.0
United Kingdom (83)	8.22
Canada (75)	7.43

Australia (59)	5.84
Germany (38)	3.76
New Zealand (34)	3.36
Other (69)	6.83
Unknown (207)	20.51

Table 2. Results of the regression model estimating day-level total pain score (N=1,009).

Conditional Random Effects		**Variance (95% CI)**	
Participant		1.09 (0.98, 1.21)	
Conditional Fixed Effects	**Rate Ratio (95% CI)**	**Log Odds (SE)**	***z*-score**
Intercept	4.26*** (3.26, 5.56)	1.45*** (0.13)	10.82
Menstrual Status	1.29*** (1.25, 1.32)	0.25*** (0.01)	20.31
Previous Day Pain	1.02*** (1.02, 1.03)	0.02*** (0.00)	29.69
Body Mass Index	1.01* (1.00, 1.02)	0.01 (0.00)	2.02
Mean Weekly Exercise Frequency	0.93* (0.89, 0.97)	-0.06** (0.02)	-2.96
Previous Day exercise	1.10* (1.05, 1.15)	0.09**(0.15)	3.88
Some College Education Level	0.87 (0.83, 1.56)	0.13 (0.15)	0.86
College or Higher Education Level	0.93 (0.66, 1.16)	-0.13 (0.14)	-0.92
Mean Weekly Exercise Frequency * Previous Day exercise	0.96** (0.95, 0.98)	-0.03** (0.01)	-3.37
Zero Inflation Terms	**Rate Ratio (95% CI)**	**Log Odds (SE)**	***z*-score**
Intercept	0.17 (0.16, 0.18)	-1.73***(0.02)	-62.96
Same Day Exercise	5.34 (5.01, 5.68)	1.67*** (0.03)	52.53

95% CI=95% Confidence Interval.*p<0.05 ** p <0.001, ***p<0.0001. Previous day pain and BMI were sample mean-centered. BMI and education level were kept as covariates in the model based on their significant associations with mean day-level pain scores (Pearson's r=0.15 for BMI and Kruskal-Wallis χ^2 = 18.061 for education level, p<0.001).

Table 3. Results of the regression model estimating pain score difference (N=1,009).

Conditional Random Effects	**Variance (95% CI)**
Participant (Intercept)	9.16 (8.28, 10.13)
Residual	26.83

Conditional Fixed Effects	**B coefficient (SE)**	**95% CI**	***z*-score**
Intercept	2.70*** (0.51)	1.68, 3.72	5.29
Menstrual Status	1.47*** (0.09)	1.28, 1.66	15.43
Previous Day Pain	-0.86*** (0.01)	-0.87, -0.85	-143.43
Body Mass Index	0.05* (0.01)	0.01, 0.10	2.86
Mean Weekly Exercise Frequency	-0.27** (0.08)	-0.44, -0.10	-3.12
Previous Day Exercise	0.92** (0.18)	0.56, 1.27	5.08
Some College Education Level	-0.84 (0.62)	-2.11, 0.42	-1.35
College or Higher Education Level	-2.07** (0.52)	-3.10, -1.03	-3.96
Mean Weekly Exercise Frequency * Previous Day Exercise	-0.14* (0.06)	-0.26, -0.01	-2.22

Zero Inflation Terms	**B coefficient**	**95% CI**	***z*-score**
Intercept	-0.91*** (0.01)	-0.93, -0.88	-63.84
Same Day Exercise	0.70*** (0.02)	0.66, 0.75	32.09

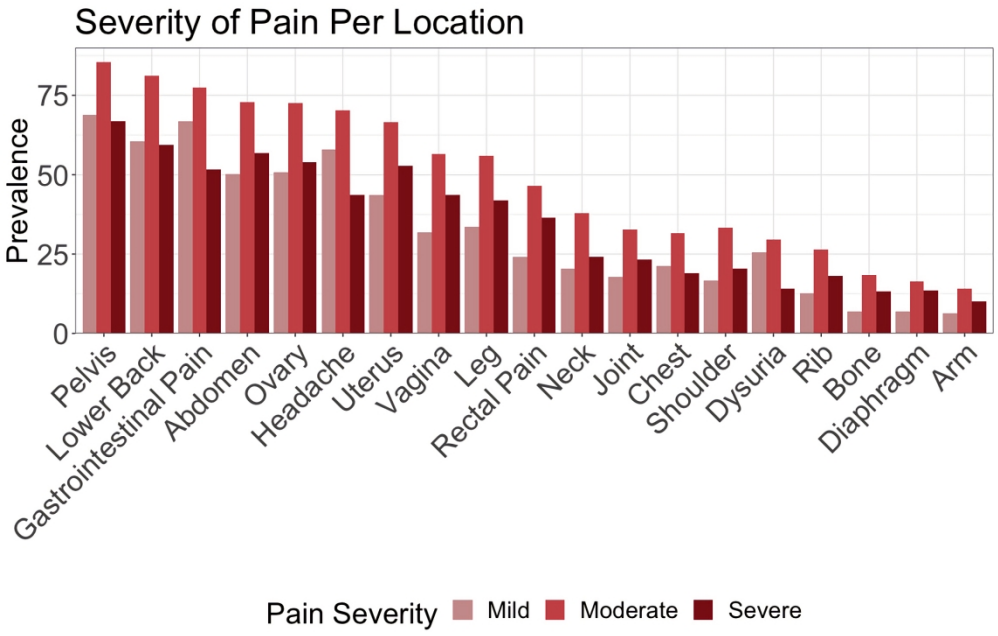
SE= Standard Error. *p<0.05 ** p <0.001, ***p<0.0001. Previous day pain and BMI were sample mean-centered. BMI and education level were kept as covariates in the model based on their significant associations with mean day-level pain scores (Pearson's r=0.15 for BMI and Kruskal-Wallis χ^2 = 18.061 for education level, p<0.001).

Figure 1. Prevalence of pain severity by location reported among participants (i.e., unique counts of body area-severity per participant). Moderate intensity was the most frequently tracked across all body areas (14.1%-85.4%).

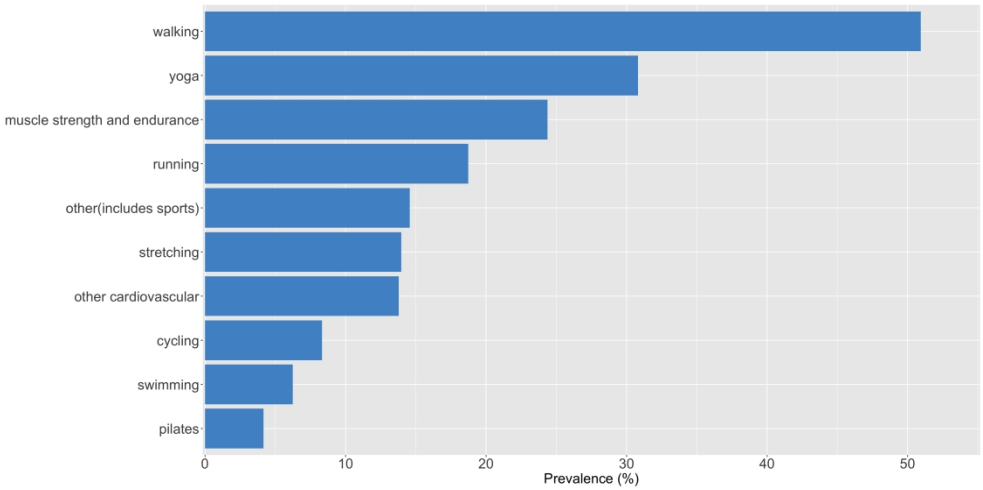
Figure 2. Prevalence of self-reported exercise modalities in the study sample. "Other cardiovascular" category include activities such as dancing, aerobics and using the elliptical machine. "Muscle strength and endurance" category includes activities such as weight lifting and calisthenics. "Other exercise" category includes sports activities such as skiing and soccer, multi-modal exercises (e.g., high intensity interval training of both cardiovascular and muscular endurance), or those that did not fit into the other categories (e.g., stabilizing or balancing exercises, wii fit or other home based fitness activities).

Panel Figure 3. Moderation of effect of previous-day exercise by habitual exercise levels (X axes). Y axes represent predicted day-level total scores (top) and differences (bottom) in pain. Shaded areas depict 95% confidence intervals. At approximately 3 times/week of regular exercise, previous day exercise starts to be associated with more favorable pain outcomes on the following day (i.e., decrease from the model predicted mean scores), adjusted for other day-level and person level factors.

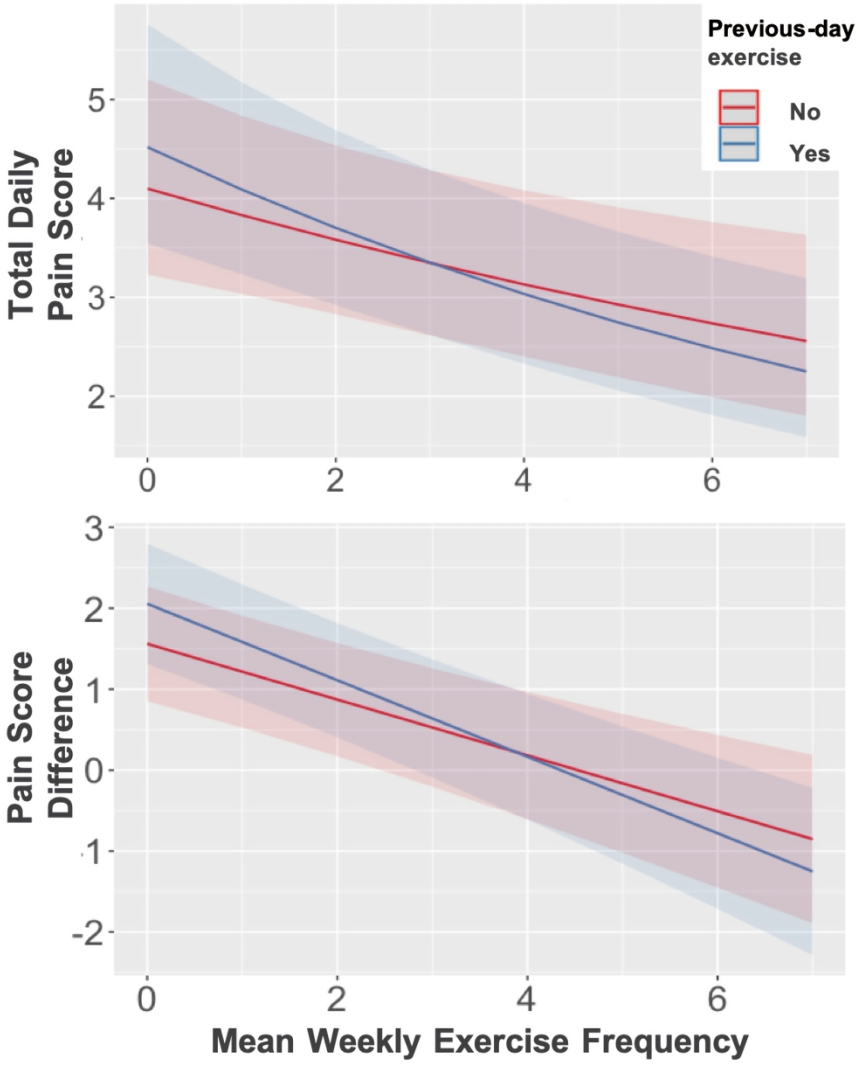
Figure 4. Plot of the random effect of the participant on total day pain scores estimated from the multilevel model (N=1,009). Y-axis represents the range of estimated average pain scores for each participant. Each black dot represents one participant's mean (i.e., random intercept), grey lines indicate 95% confidence intervals. Distribution of points across the x-axis indicate large variability across individuals (i.e., between-group variance), and the grey lines indicate the within-person variability in daily scores over time.



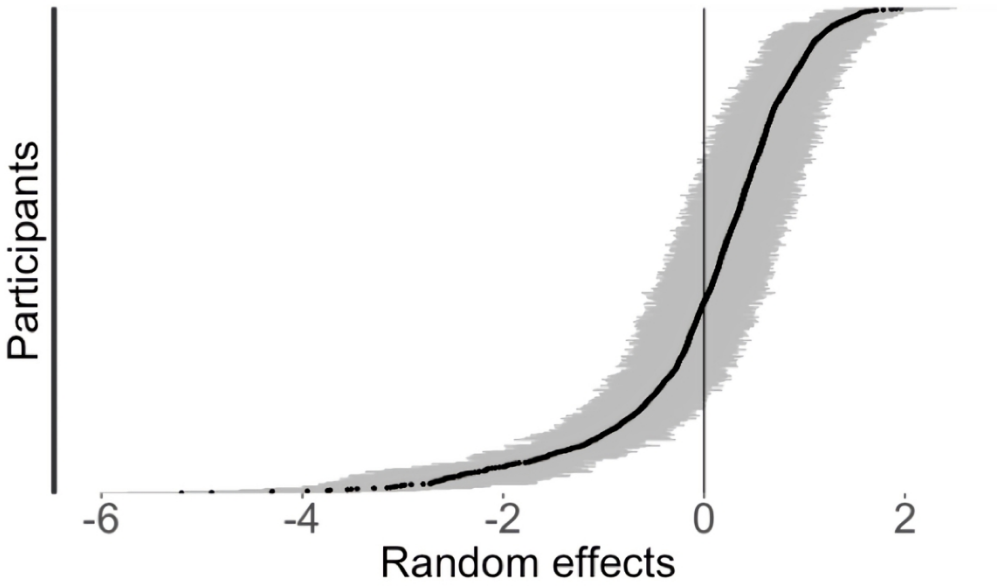
776x493mm (72 x 72 DPI)



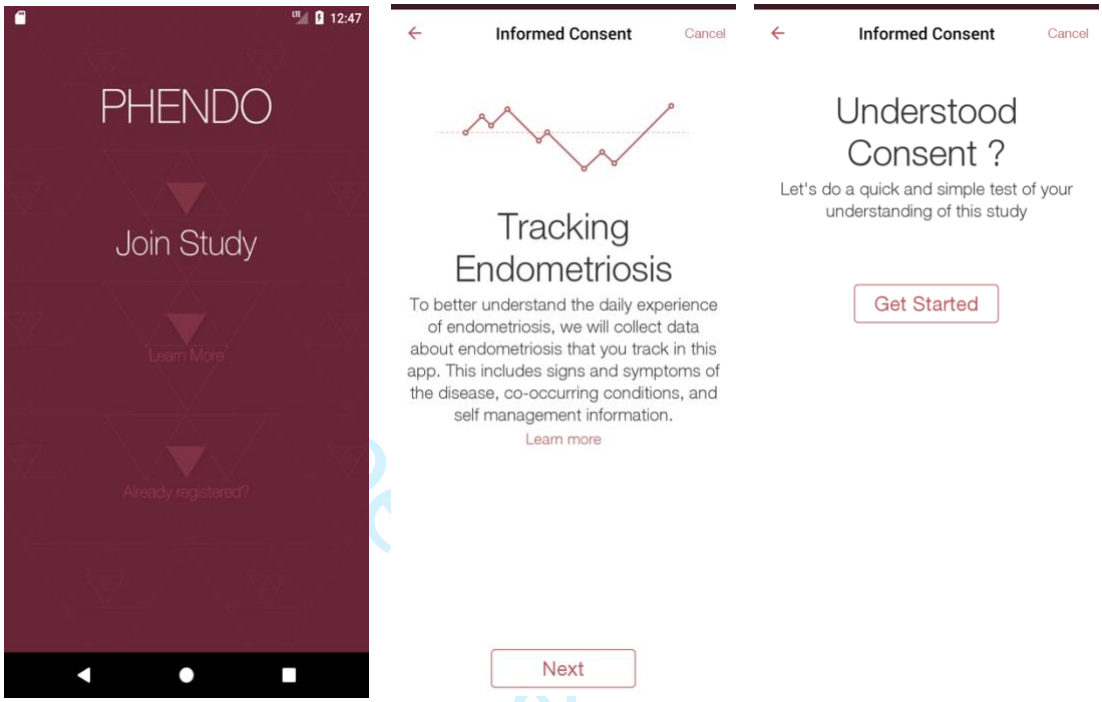
1411x705mm (72 x 72 DPI)



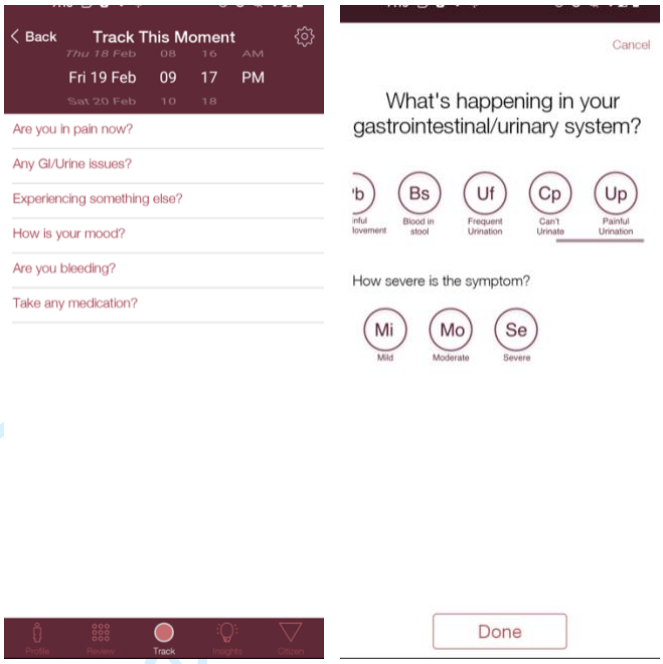
694x750mm (72 x 72 DPI)



571x339mm (72 x 72 DPI)



Supplementary Figure 1. Screenshots of the Phendo Registration/download page (left), beginning of the informed consent obtainment (middle), and self-quiz to verify consent (right).



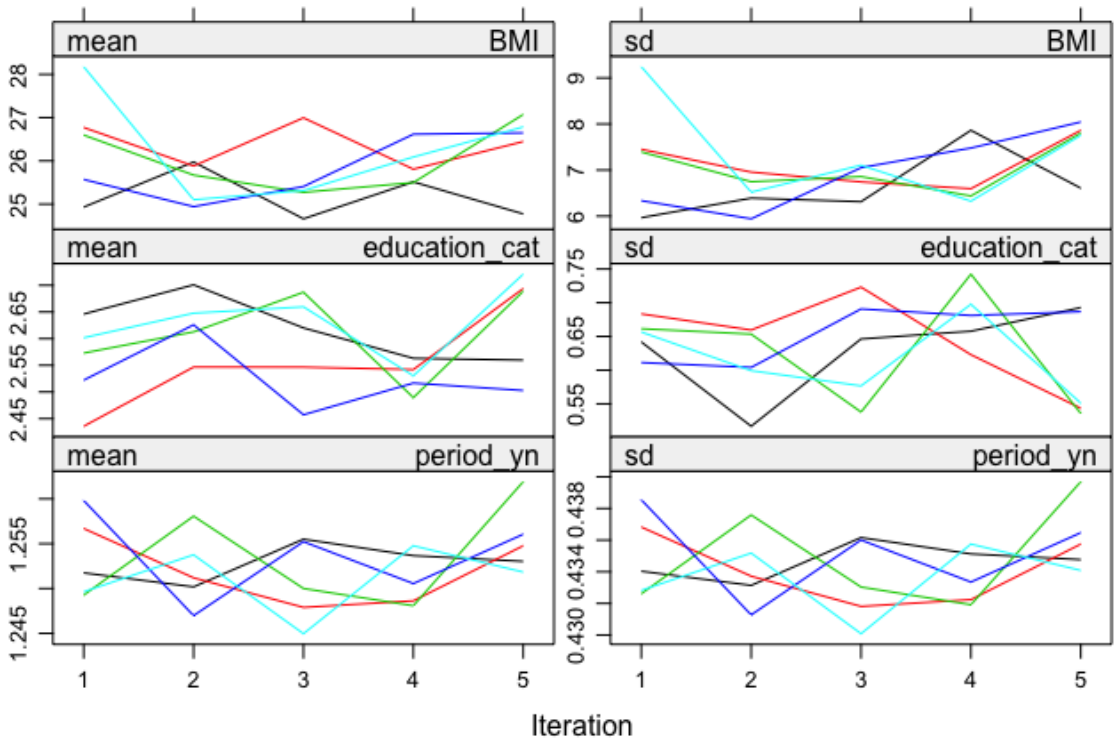
Supplementary Figure 2. Screenshots of Phendo's momentary tracking tab (left) and an example individual symptom and severity tracking (right).

Supplementary File 1. Missing Data Imputations.

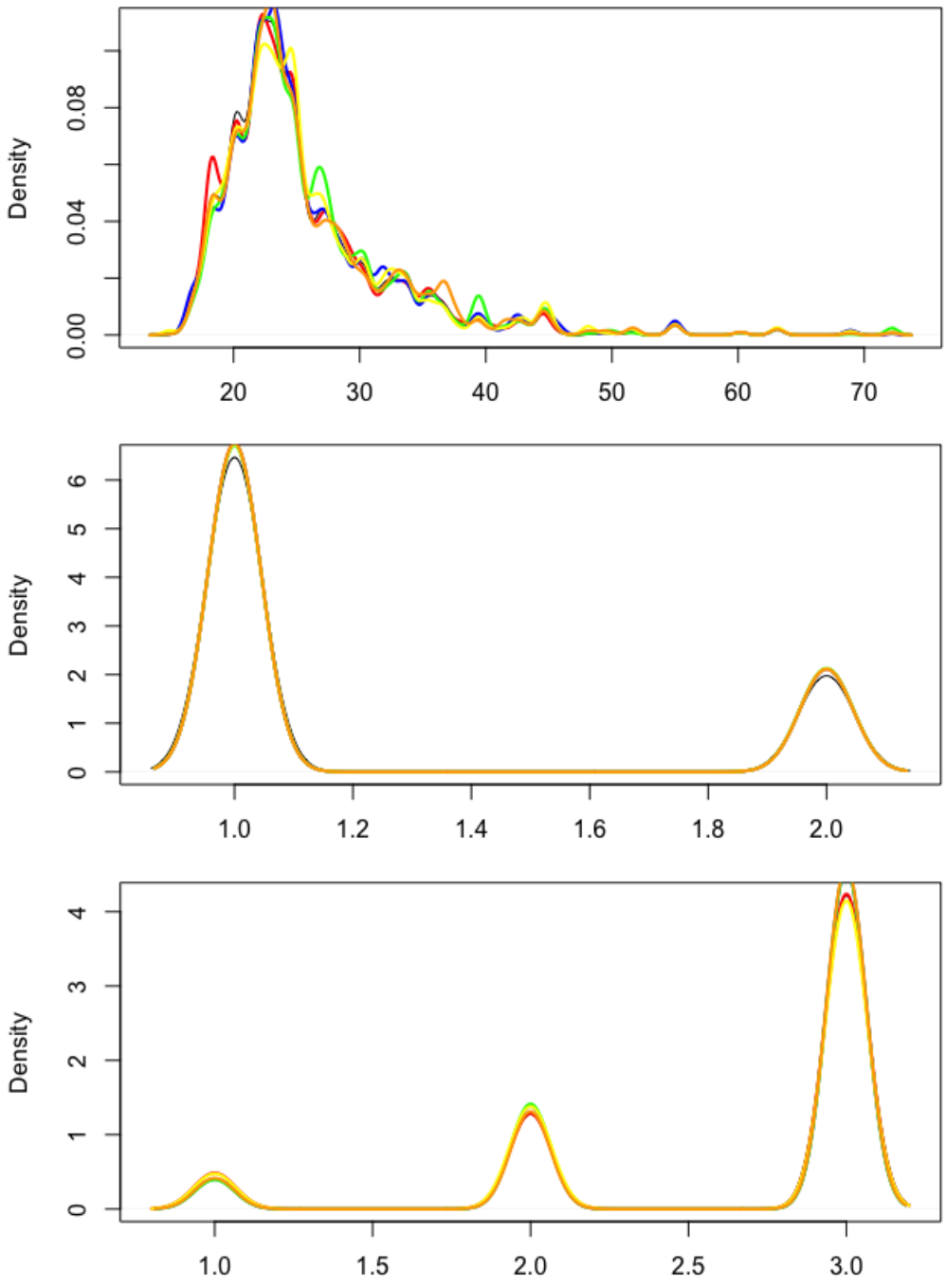
Because Phendo is an observational research app and participants are free to track (or not track) any given item as they so wish and do not receive prompts from the research team to track any given item at a certain time, missingness in the data occurs due to a variety of possible reasons that are difficult to distinguish. For example, a period not tracked for a day could mean that the participant did not have a period, or they chose not to track, or did not use the app at all that day. To circumvent this issue, we took several measures. First, we limited data to days for which the participant tracked their pain, exercise and menstrual status at least once, as a proxy for app use. Next, we assigned a score of zero for pain on days where the participant had tracked exercise or menstrual status but not pain. This approach is motivated by 2 reasons. First, the nature of the pain question in Phendo (i.e., "Where is the pain?", "How severe is the pain?") assumes the participants to track when they feel pain and therefore a "No Pain" response is neither available in the app nor would make sense. Second, multiple imputation methods impute such that the resulting imputations are limited to the observed values and distributions. Thus by default it would omit the possibility of a zero in the resultant pain score distribution, which increases risk of overestimation of the scores in the sample.

BMI (calculated from participant reported height and weight) and education level were missing for 22% and 19% of the participants, respectively, and menstrual status was missing (i.e., not tracked) 22% of the time in the dataset. We imputed these 3 variables using multivariate imputations by chained equations [83] according to the heteroscedastic linear two-level (i.e., hierarchical where, participant is the clustering variable) structure of the data following published guidelines on multilevel multiple imputation methods. [83, 84, 127, 128] We used

two-level predictive mean matching for BMI and education level, which is a semi-parametric imputation method that limits imputations to the observed values and can preserve non-linear relations in the observed data, therefore the imputations do not deviate from the observed distribution[129] and two-level logistic regression for imputing menstrual status, using the rest of the dataset as the predictors. As per published recommendations,[83, 84] we also included the raw pain variable (i.e., with the missing values) as a predictor, to account for the possibility of an association between the missingness pattern of pain to these imputed variables. To assess the plausibility of the imputations and any significant deviance from the structure of the raw, non-imputed data, we inspected the imputation convergence plots, distributions of the imputed variables which are provided in Supplementary Figures 3 and 4.



Supplementary Figure 3. Convergence plots for the 3 imputed variables (BMI, top; education, middle; menstrual status, bottom) with means on the left and standard deviations on the right side of the panel. Plots indicate healthy convergence based on lack significant trend and the streams mingling well right from the start throughout the 5 iterations (x-axis).



Supplementary Figure 4. Density plots of the marginal distributions of BMI (top), menstrual status (middle), and education category (bottom) of raw, non-imputed data and 5 iterations of the imputed data. Close super-imposition of the curves indicate that the imputed data distributions match those of raw data.

Supplementary File 2. Imputation Model diagnostics.

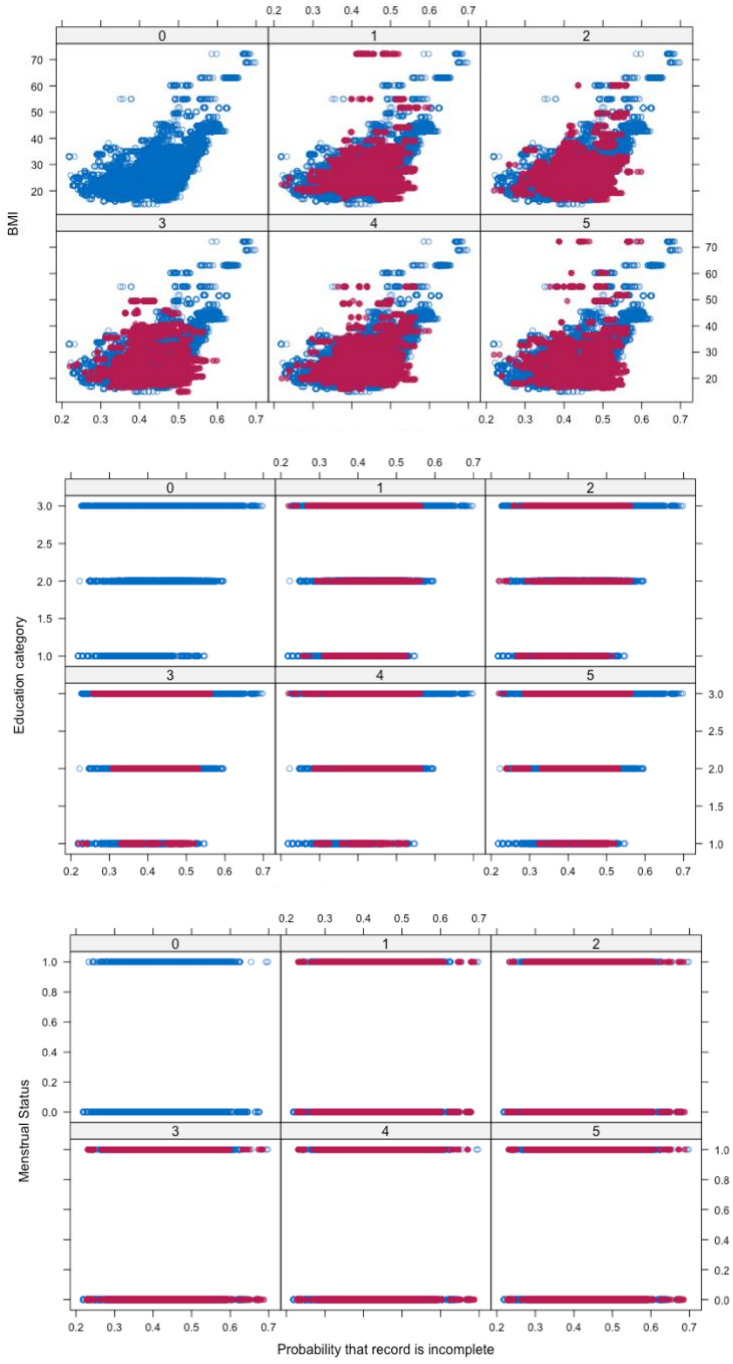
Appropriateness and plausibility of the estimates from imputed models were inspected following published guidelines. First, we used measures of missing data information to assess pooled estimate variances. The fraction of missing information (λ) is interpreted as the proportion of variation in the parameter of interest due to the missing data. The relative increase in variance due to nonresponse (r) is interpreted as the proportional increase in the sampling variance of the parameter of interest that is due to the missing data. Values of λ over 0.5 indicate that the influence of the imputation model on the results is larger than that of the complete-data model, suggesting potential problems in the imputations. Supplementary Table 1 provides results of these variance estimates, indicating satisfactory imputation and model fit.

Supplementary Table 1. Measures of Missing data information

	Total Pain Score		**Difference in Pain**	
Conditional Fixed Effects	**λ**	***r***	**λ**	***r***
Intercept	0.21	0.27	0.23	0.31
Menstrual Status	0.13	0.15	0.19	0.23
Previous Day Pain	0.01	0.01	0.00	0.00
Body Mass Index	0.13	0.15	0.23	0.31
Mean Weekly Exercise Frequency	0.00	0.00	0.01	0.01
Previous Day exercise	0.01	0.01	0.00	0.00
Some College Education Level	0.26	0.36	0.35	0.55
College or Higher Education Level	0.23	0.31	0.21	0.28
Mean Weekly Exercise Frequency * Previous Day exercise	0.00	0.00	0.00	0.00
Zero Inflation Terms				
Intercept	0.00	0.00	0.00	0.00
Same Day Exercise	0.00	0.00	0.00	0.00

Next, we inspected propensity scores, which is a more recent and increasingly accepted method for inspecting the suitability of data imputation.[84, 85, 130] The goal is to compare the distributions of observed and imputed data conditional on the missingness probability. Under the missing at random (MAR) assumption, the conditional distributions of the observed and missing data should be similar if the assumed model for creating multiple imputations has a good fit. To do this, we first estimate the probability of each record being incomplete (i.e., "response propensity") in the presence of missing data by conditioning on the response indicators as well as the observed covariates. The probabilities are then averaged over the imputed datasets to obtain stability. Supplementary Figure 3 plots BMI, education category and menstrual status against the propensity score in each dataset. The distributions of the blue

and red points are match up well without significant discrepancies (e.g., mismatch in patterns, imputed data systematically shifted toward one side of the axis).



Supplementary Figure 5. BMI (top), education category (middle) and menstrual status (bottom) plotted against the propensity score in each dataset (0=observed, 1-5=imputed). The distributions of the blue (observed) and red (imputed) points are follow similar patterns.

Supplementary Table 2. Post-hoc analyses with endometriosis diagnosis included as a covariate. Conditional model results of the negative binomial model estimation of day-level total pain score (N=608).

Random Effects	**Variance (95% CI)**
Participant (Intercept)	1.10 (0.99, 1.22)

Fixed Effects	**Log Odds (SE)**	**z-score**
Intercept	1.37*** (0.12)	10.97
Menstrual Status	0.25*** (0.01)	21.40
Previous day Pain	0.02*** (0.01)	21.40
Body Mass Index	0.01* (0.004)	2.81
Mean weekly Exercise Frequency	-0.06** (0.02)	-3.01
Previous day exercise	0.09** (0.02)	3.85
Clinician diagnosis of endometriosis	-0.07 (0.10)	0.01
Self-diagnosis of endometriosis	-0.11 (0.11)	-1.01
Some college education level	0.22 (0.13)	-1.63
College or higher education level	-0.01 (0.12)	-0.12
Mean weekly Exercise Frequency*Previous day exercise	-0.03*** (0.01)	-3.42

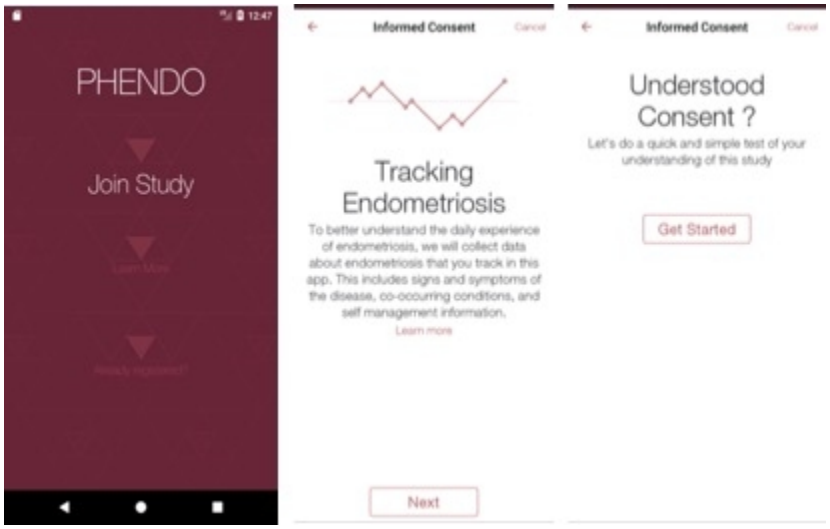
SE=Standard Error. *p=0.001, ** p <0.001, ***p<0.0001. B coefficients are rate ratios. BMI =Body Mass Index. BMI and previous day pain were group mean centered.

Supplementary Table 3. Post-hoc analyses with endometriosis diagnosis included as a covariate. Conditional model results of the regression model estimation of pain score difference (N=1009).

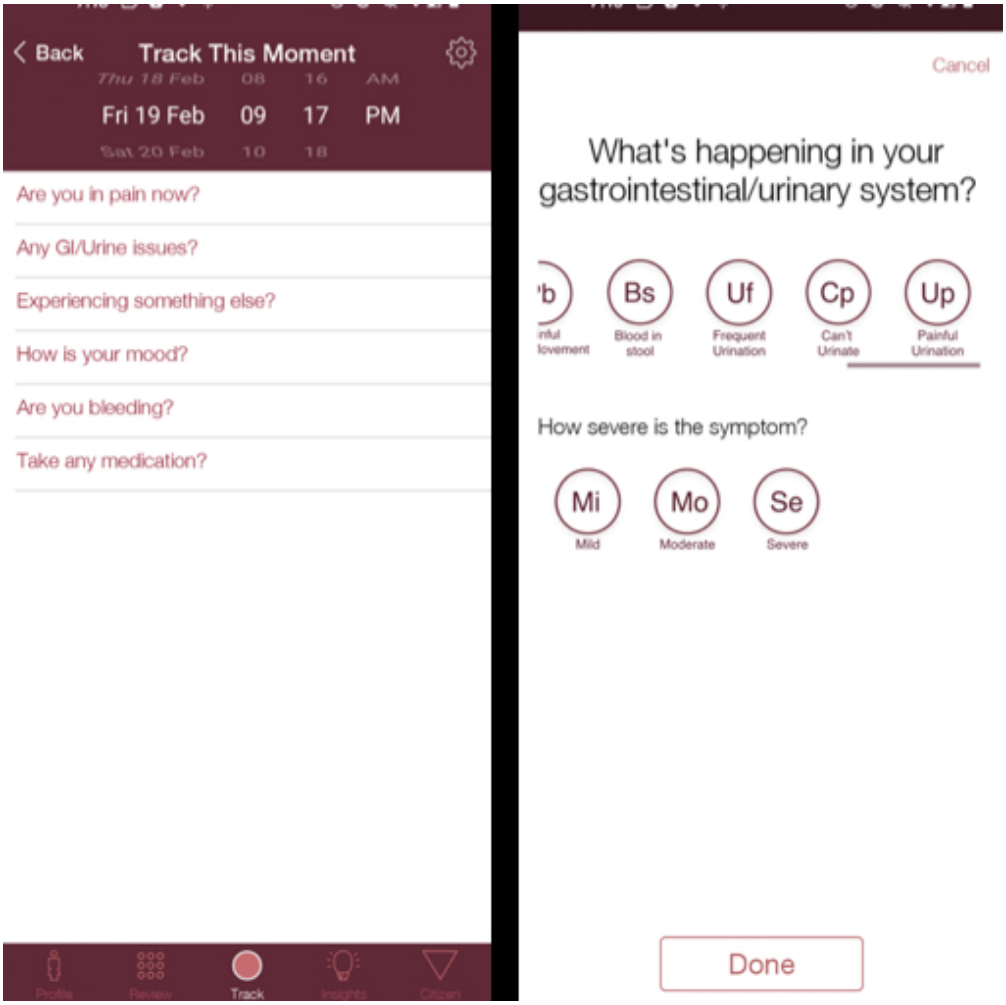
Conditional Random Effects	**Variance (95% CI)**
Participant (Intercept)	13.34 (12.09, 14.93)

Fixed Effects	**B coefficient (SE)**	**z-score**
Intercept	2.45*** (0.46)	5.22
Menstrual status	1.46*** (0.08)	16.98
Previous day pain	-0.86*** (0.01)	-144.11
Body mass index	0.07* (0.01)	4.47
Mean weekly exercise frequency	-0.27** (0.09)	-3.03
Previous day exercise	0.92*** (0.18)	5.13
Clinician diagnosis of endometriosis	-0.05 (0.32)	-0.16
Self-diagnosis of endometriosis	-0.45 (0.43)	-1.29
Some college education level	-0.30 (0.51)	-0.58
College or higher education level	-1.72** (0.47)	-3.67
Mean weekly exercise frequency*Previous day exercise	-0.14* (0.06)	-2.31

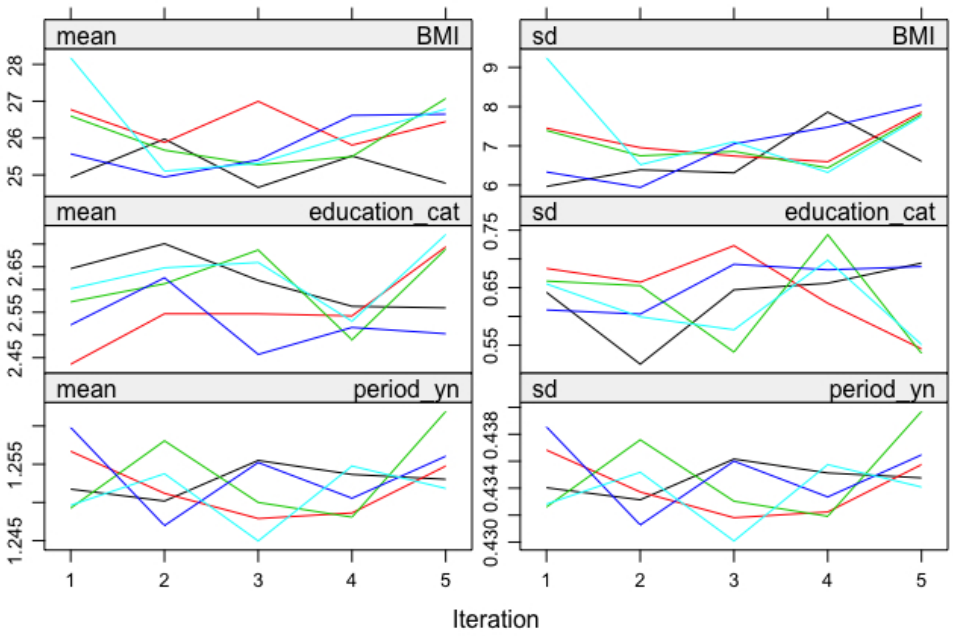
SE=Standard Error. *p<0.05, ** p <0.01, ***p<0.0001. Body Mass Index and previous day pain were group mean centered.



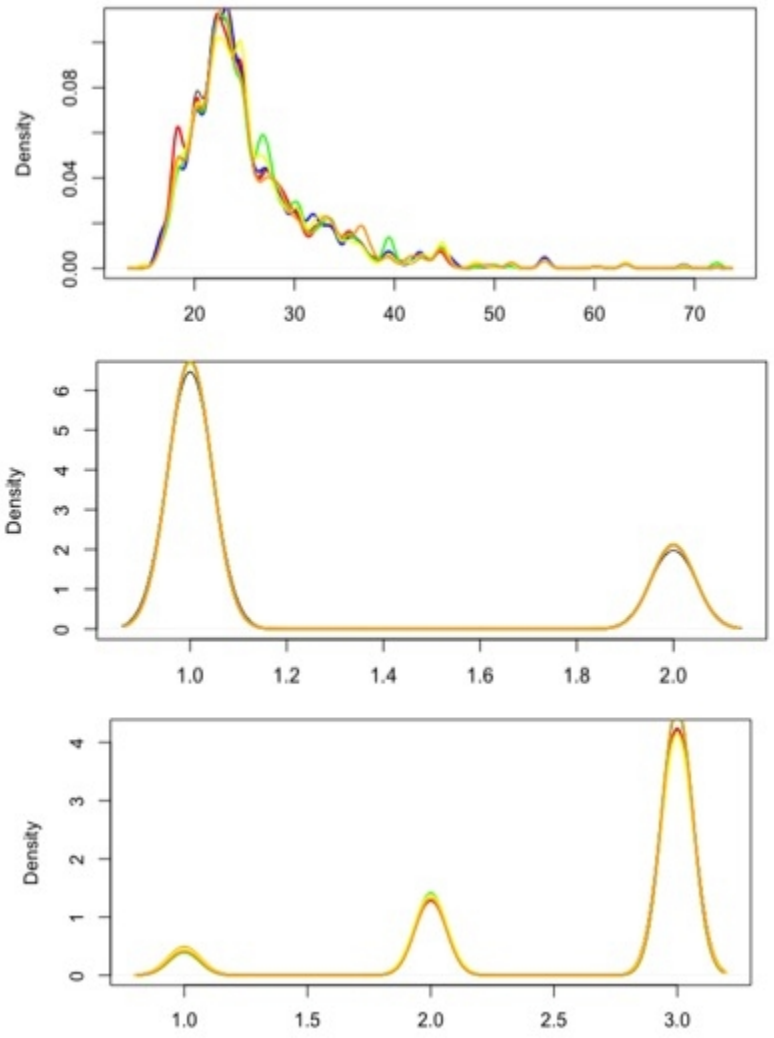
34x22mm (300 x 300 DPI)



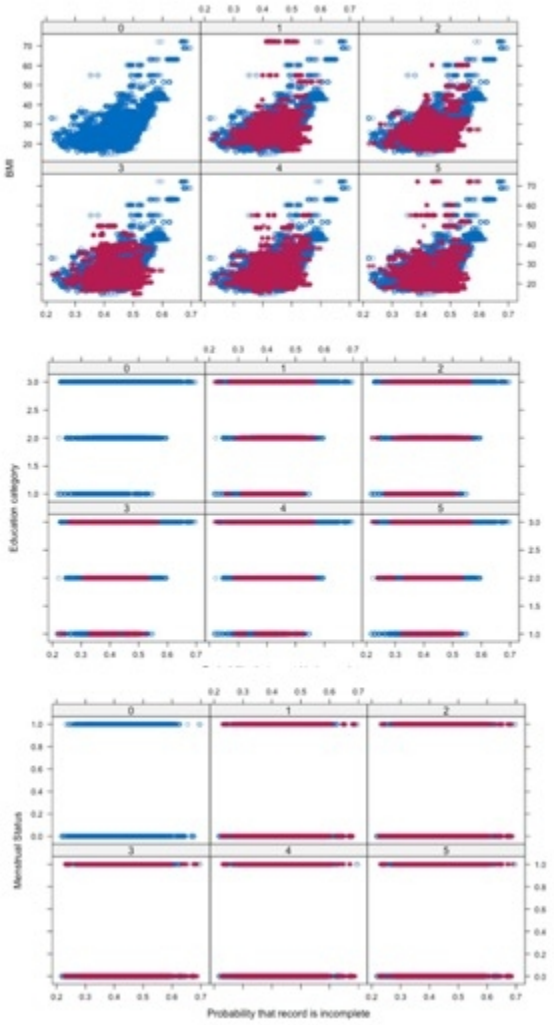
109x109mm (118 x 118 DPI)



142x99mm (118 x 118 DPI)



34x45mm (300 x 300 DPI)



23x44mm (300 x 300 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page # where this item is located:
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3-5
Objectives	3	State specific objectives, including any prespecified hypotheses	6
Methods			
Study design	4	Present key elements of study design early in the paper	6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-7
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	6-7
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed	N/A

		Case-control study—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-10
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*7-10*
Bias	9	Describe any efforts to address potential sources of bias	6-8
Study size	10	Explain how the study size was arrived at	6-7

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8-10
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	11-12
		(*b*) Describe any methods used to examine subgroups and interactions	11-12
		(*c*) Explain how missing data were addressed	12, Supplemental Files 1-2
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	11-12
		(*e*) Describe any sensitivity analyses	12, and Supplemental Tables
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	13
		(b) Give reasons for non-participation at each stage	6, 13
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	13, and Table 1
		(b) Indicate number of participants with missing data for each variable of interest	13, Table 1
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	N/A
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	*N/A*
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	*N/A*
		Cross-sectional study—Report numbers of outcome events or summary measures	13,14
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	14, Tables 2 and 3

	(*b*) Report category boundaries when continuous variables were categorized	14-15
	(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	14-15, and Supplemental Tables
Discussion			
Key results	18	Summarise key results with reference to study objectives	16
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	21-22
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	16-19
Generalisability	21	Discuss the generalisability (external validity) of the study results	21,22
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	23

BMJ Open

An Observational Study of Exercise and Pain Patterns in Endometriosis via Mobile Self-Tracking

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-059280.R1
Article Type:	Original research
Date Submitted by the Author:	26-Apr-2022
Complete List of Authors:	Ensari, Ipek; Columbia University, Data Science Institute Lipsky-Gorman, Sharon; Columbia University, Department of Biomedical Informatics Horan, Emma; Columbia University, Department of Biomedical Informatics Bakken, Suzanne; Columbia University, School of Nursing Elhadad, Noemie; Columbia University, Department of Biomedical Informatics
Primary Subject Heading:	Health informatics
Secondary Subject Heading:	Public health
Keywords:	PAIN MANAGEMENT, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS, PREVENTIVE MEDICINE, EPIDEMIOLOGY, COMPLEMENTARY MEDICINE

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

An Observational Study of Exercise and Pain Patterns in Endometriosis via Mobile Self-Tracking

Running title: Daily exercise and pain patterns in endometriosis

Ipek Ensari[1], PhD, Sharon Lipsky-Gorman[2], MA, Emma Horan[2], BS, Suzanne R. Bakken[3], PhD, FAAN, Noémie Elhadad[2], PhD

[1]Data Science Institute, Columbia University, [2]Department of Biomedical Informatics, Columbia University Irving Medical Center, [3]Columbia University School of Nursing, New York, NY

Author of Correspondence:
Ipek Ensari
Columbia University
Data Science Institute
475 Riverside Dr, Room 320,
New York, NY 10115
Email:ie2145@columbia.edu

Word count: 4,490

Abstract

Objectives: This study investigates the association of daily physical exercise with pain symptoms in endometriosis. We also examined whether an individual's typical weekly (i.e., habitual) exercise frequency influences (i.e., moderates) the relationship between their pain symptoms on a given day (day *t*) and previous-day (day *t-1*) exercise. **Participants:** The sample included 90,382 days of data from 1,009 participants (~85% non-Hispanic white) living with endometriosis across 38 countries. **Study Design:** This was an observational, retrospective study conducted using data from a research mobile app (Phendo) designed for collecting self-reported data on symptoms and self-management of endometriosis. **Primary Outcome Measures:** The two primary outcomes were the composite day-level pain score that includes pain intensity and location, and the change in this score from previous day (Δ-score). We applied generalized linear mixed-level models to examine the effect of previous-day exercise and habitual exercise frequency on these outcomes. We included an interaction term between the 2 predictors to assess the moderation effect, and adjusted for previous-day pain, menstrual status, education level, and body mass index. **Results**: The association of previous-day (day *t-1*) exercise to pain symptoms on day *t* was moderated by habitual exercise frequency, independent of covariates (Rate ratio=0.96, 95% CI=0.95, 0.98, p=0.0007 for day-level pain score, B=-0.14, 95%CI=-0.26, -0.016, p=0.026 for Δ-score). Those who regularly engaged in exercise at least 3 times per week were more likely to experience favorable pain outcomes after having a bout of exercise on the previous day. **Conclusions:** Regular exercise might influence the day-level (i.e., short-term) association of pain symptoms to exercise. These findings can inform exercise recommendations for endometriosis pain management, especially for those who are at greater risk for lack of regular exercise due to acute exacerbations in their pain after exercise.

Strengths and limitations of this study

- This study leverages data from a research mobile app (Phendo) designed for collecting self-reported data on symptoms and self-management of endometriosis.
- Daily exercise and pain symptom patterns in endometriosis is investigated under ecologically-valid conditions.
- The participant sample (N=1,009) represents 38 countries, ages across the reproductive life span, and various person-level characteristics.
- The study is limited by self-reported data collection by somewhat consistent trackers and lacks details on duration or intensity of exercise to evaluate as potential moderators.
- Participants consisted of mostly white, non-Hispanic individuals, therefore results might not be generalizable to other demographic groups.

INTRODUCTION

Exercise, a subset of physical activity (PA) that is planned, structured, repetitive, and intended to improve or maintain physical fitness, is an important component of effective pain management (i.e., reduction and prevention of pain symptoms).[1, 2] Both acute (i.e., single bout/session) and chronic (i.e., repeated bouts/sessions over time) exercise training have been demonstrated to improve numerous pain-related conditions. [1, 3-7] However, pain-related responses to exercise appear to be variable in populations with chronic pain conditions.[8] Similarly, exacerbation of pain with exercise could pose a barrier to regular exercise in such individuals, thus increasing resistance to exercising, which in return can worsen pain, related disability, and risk for co-morbidities.[9-11] Investigation into the naturally-occurring pattern of pain symptoms associated with exercise behavior can help inform the design of exercise-based therapies for targeting disease-related pain symptoms.

Individuals with endometriosis may benefit from such an investigations for several reasons.[12-14] Endometriosis is a systemic, estrogen-dependent inflammatory condition characterized primarily by chronic pelvic and abdominal pain, pain with sexual intercourse, and infertility.[15, 16] It significantly impacts daily function and quality of life (QoL)[17, 18], contributing to a productivity loss of 6.3 hours/week[19] and an estimated $69.4 billion in excess health expenditures annually in the United States.[20] Existing medical and hormonal therapies have limited efficacy on pain management, often confounded by side effects.[21] Opioids and other analgesics are commonly prescribed for long-term use,[22, 23] despite treatment guidelines recommending use of nonpharmacologic therapies including PA.[24] Consequently, there is a critical need to identify alternative approaches for endometriosis pain management.

One such approach is exercise, based on various mechanisms proposed in the literature[25] that might pertain to endometriosis. These include regulation of the serotonergic and opioid receptors,[26] reduction of inflammatory markers associated with pain,[27, 28] and effect of exercise on nerve growth factor expression that is associated with the painful endometriosis lesions.[29, 30] Exercise can increase pain management self-efficacy, which is associated with improved pain outcomes and QoL for individuals with chronic pain.[31] While the evidence on exercise for pain management is promising [4, 32, 33], existing data are scarce, cross-sectional, and indicate variable effects on pain outcomes.[33-37] Despite these limitations, previous reports of exercise-induced adaptations to pain stimuli through increased pain threshold suggest that the regularity with which an individual engages in exercise over the long term (i.e., habitual exercise frequency) might influence (i.e., moderate) the relationship between their day-level exercise and pain symptoms.[38, 39] Among regular exercisers, pain-related activation has been demonstrated in the brain's descending antinociceptive pathway, with corresponding reductions in self-reported pain after acute bouts of at least moderate intensity exercise.[40] Moreover, studies report that habitual exercise frequency moderates a variety of self-reported outcomes (e.g., mood, anxiety, fatigue) in response to acute exercise.[41-43] While these findings are promising, their generalizability are limited by sample characteristics, laboratory-based experimental pain stimuli and exercise manipulations, and brief measurement duration of up to several hours. Thus, further investigation is needed to examine the relationship between pain symptoms and exercise behavior with a representative sample, under ecologically valid conditions, while accounting for possible between-individual variability and temporal lags in the outcome that extend beyond several hours.

Accordingly, this study examines the naturally-occurring daily patterns of pain symptoms and exercise behavior in endometriosis. We leverage mobile self-tracking, a particularly useful approach for capturing ecologically valid profiles of the dynamic temporal fluctuations and between-individual variability in pain over time.[44] We primarily aim to delineate the degree to which an individual's typical weekly exercise frequency (i.e., habitual exercise) influences (i.e., moderates) the association of their pain symptoms on a given day (day *t*) to their previous-day (day *t-1*) exercise behavior (i.e., lagged-day effects). Given the previously documented variable course of pain symptomology in endometriosis,[45] we also delineate the variability in day-to-day pain experiences within these analyses.

MATERIALS AND METHODS

Study Design

Study design and protocols were approved by the Columbia University Irving Medical Center (CUIMC) Institutional Review Board (#AAAQ9812). This study was conducted with retrospective data collected through an observational research mobile app "Phendo". Phendo was designed and developed for self-tracking endometriosis symptoms and its management. It is available for iOS[1] and Android[2] in App stores for free.

Study Sample and Inclusion Criteria

The study sample comprised Phendo users with a self-reported surgery-, clinician-, or suspected diagnosis of endometriosis and self-tracked exercise and pain data between November 2016 and April 2020. All participants regardless of diagnosis type are provided the same set of measures in the App. In a previous study, the endometriosis phenotype (i.e., characterization) obtained using Phendo data was demonstrated to be consistent with both the characterization of

the disease in the literature based on standard clinical surveys and with clinician (i.e., human expert) evaluations.[46] We a priori decided to include all participants who selected one of the three affirmative responses in the present analyses, excluding those who indicated not having endometriosis. Out of the initial eligible pool of 9,792 Phendo users with reported endometriosis, 7,949 had at least one day of tracking of the variables of interest for the study. Of these, 1,009 users had sufficient amount of data on pain and exercise for analysis (See *Data Analysis*) and thus were included in the study.

Recruitment and Informed Consent

Study participants were passively recruited through one of the App stores, engagement on study social media sites, or word-of-mouth. Upon downloading Phendo, all potential users go through an informed consent and enrollment process before tracking any data. First, they are provided with an explanation of the App, its overall purpose and link to its website (citizenendo.org) which includes etailed information and instructional videos for using the App. Participants complete a brief "verify your understanding" quiz to ensure their comprehension of how their data might be used for research purposes, anonymity and confidentiality (See Supplementary Figures 1-2 for example screenshots). This is followed by formal electronic informed consent (and assent for individuals 13-18 years old), a copy of which is sent to the participant. Once enrolled, users are instructed to track daily, but they are free to track as much or as sporadically as they wish, and they do not receive any prompts or requests to track a specific variable from the research team. Findings from a previous study evaluating recruitment and retention patterns within Phendo and seven other similar self-tracking apps indicated that Phendo's engagement was similar to standard engagement patterns in research smartphone

apps.[47] Participants in the current study did not receive financial compensation for their tracking activities.

Patient and Public Involvement

Measures in Phendo were developed using patient-centered participatory design, through qualitative (focus groups, interviews) and quantitative research (surveys, coded content analysis) with participants with endometriosis, described in detail elsewhere.[48, 49] Studies suggest that this technique for developing patient-reported outcome measures enhances content validity and relevance of the measure to the target population, thus providing a more comprehensive and accurate representation of the disease under study.[50-53]

Study Measures

Day-level Pain. We assessed day-level pain through multiple items within Phendo: 1. "Are you in pain now? Where is the pain?", 2. "Any gastrointestinal or urinary issues?" (painful urination (dysuria), painful bowel movement (dyschezia)). Similar to pain documentation in clinical records and other measures such as the McGill Pain Scale,[54] Phendo users can select location from all areas of the body (20 available choices, as well as right/left and upper/middle/lower specification), and can be mapped onto a visual, analogous to the McGill Pain Scale. Phendo users rate severity for each affirmative response on a 3-point categorical scale (mild, moderate, or severe), analogous to other commonly used pain rating scales in the literature.[55, 56] This categorization has been used for standardization and comparisons across different pain measures, and demonstrated superior ability to capture the nonlinear relationship between reported pain severity and interference with activity than use of numbers.[57, 58]

While mHealth studies have examined the validity, utility and specificity for various pain conditions[58-60] of their pain measurement approaches, a standard "all-in-one" single outcome

that captures the multi-dimensional pain experience across different populations remains to be established.[61, 62] Thus, composite pain computations have been proposed.[63] We computed a heuristic, composite day-level pain score to capture participants' conceptualization of their pain experience by summing the severity scores reported for each body area (e.g., moderate pain in abdomen, mild pains in chest and leg would yield *2+1+1=4* as the total score).[45] To account for and circumvent any potential pain rumination/catastrophizing [62, 64] and varying tracking habits among participants, the score was computed based on the unique reports of area-severity pairs per day for each participant (e.g., if a participant tracked mild abdominal pain three times in a day, this abdomen-mild pair is counted toward the daily pain score only once). This score was the foundation of two study outcome variables: 1) total day-level pain score, and 2) difference in day-level pain score from previous day to the next (i.e., *t*-(*t*-1)). The latter captures additional nuances in the data, enabling analyses to distinguish between participants with overall high day-level pain scores over time and experience a post-exercise reduction in pain versus those with low pain scores and who not experience a post-exercise reduction in pain. In the current study sample, the composite pain scores were moderately correlated with scores from other standard pain measures (e.g., $r=0.36$, $p<0.0001$ with the Pelvic-Abdominal Pain Visual Analog Scale (VAS); $r=-0.46$, $p<0.0001$ with Medical Outcomes Study 36-item Health Survey (SF-36) Bodily Pain subscale).

Day-level and habitual exercise. Phendo users track their daily exercise through responding to a root question "Did you exercise today? (Yes/No)". Upon selecting a "Yes", users can further customize their entry within this item by adding exercise details through unrestricted free-text responses. We used responses to the root item to compute day-level and mean weekly exercise frequency (i.e., habitual exercise) for each participant. We calculated the latter by

summing the number of exercise reports tracked per week across the range of days of data and then dividing this number by the total number of weeks of data. We used free-text responses to categorize exercises by modality and to validate that entries were exercise-related. Any non-exercise activity (e.g., sleep, meditate, sitting, socialize) was recoded as a no exercise in the analytic data set. This day-level exercise assessment aims to increase ecological validity[50, 65] and reduce the likelihood of low test-retest reliability and inaccuracy due to recall bias.[66] Similar mHealth measures of daily PA and exercise have been used by others[67-69] who reported concordance with accelerometer-based measures,[70] and higher correlations than self-report methods with accelerometer measures.[67, 68] We evaluated the validity of the scores from the Phendo exercise item through a series of analyses with the study sample [71] Results supported its concurrency with other self-reported recall-based measures (i.e., kendall's τ=0.256, $p<0.001$ with Exercise Vital Sign[72] and τ=0.294, p=0.001 with accelerometers; B=18.73, p=0.039 in association to the Nurses' Health Study II Weekly Exercise Scale[73] scores).

Standard Pain and Exercise Measures. To allow comparisons of the study sample with others in the literature, we report sample summary scores from the following components of the World Endometriosis Research Foundation (WERF) Endometriosis Patient Questionnaire (EPQ-S)[74, 75]: 1) The 2-item Bodily Pain subscale of the SF-36,[76] 2) Pelvic-abdominal Pain VAS ("*Please rate how severe your general pelvic/lower abdominal pain was at its worst in the last 3 months using the pain scale below where 0=no pain and 10=worst imaginable pain*."), and 3) The 8-item Nurses' Health Study II Weekly Physical Activity Scale (NHS-II) [73]. It measures self-reported weekly durations of major exercise modalities (i.e., walking, running, lap swimming, jogging, bicycling, tennis, calisthenics, other aerobic recreation) in a typical week in the past 12 months. These durations can further be multiplied by their metabolic equivalents

(METs) based on the Compendium of PA [77] and summed to obtain the total weekly exercise-related energy expenditure (EE). We report both the total weekly minutes and EE for the sample.

Data Analysis

Sample Characteristics. We characterize the study sample through frequencies (%) and means (standard deviation; SD) of demographics, self-reported pain medication use habits, and scores on the standard pain and exercise measures for those who completed the surveys. We characterize pain symptomology in the sample by describing the prevalence of self-tracked pain severities by each body area.

Associations of pain symptoms with exercise behavior. Using generalized linear mixed models (GLMMs), we separately estimated day-level total pain score and pain score difference as primary outcomes. Both outcomes were regressed on previous-day (day *t-1*) exercise and mean weekly exercise frequency to estimate the slope of mean pain level on day *t* and change in pain. We included an interaction term between the 2 predictors to assess the moderation of the day-level association by each individual's mean weekly exercise frequency. We included participant as a random effect to account for between-person variability in daily pain by estimating a separate intercept for each participant. Models were further adjusted for menstrual status (binary: yes/no), previous-day (i.e., day *t-1*) pain, body mass index (BMI) and education level. Race/ethnicity and age were not significantly associated with average daily pain reports ($F=1.68$, $p=0.14$ for race/ethnicity; $r=-0.148$, $p=0.07$ for age), and age was further significantly associated with education level (Kruskal-Wallis $X^2=64.948$, $p<0.0001$). To avoid redundancy and multicollinearity, race/ethnicity and age were not included as model covariates.

Model Specification. We specified a zero-inflated negative binomial (ZINB) distribution when modeling the total pain outcome, as it has been demonstrated to provide the best fit for outcomes with over-dispersion and zero-inflation (i.e., zeros due to both sampling and missingness) [78-80]. Missing values in the BMI (22%), education level (19%) and menstrual status (22%) were imputed as described in Supplementary File 1 and checked for appropriateness based on convergence and marginal distributions following guidelines [81-83] (See Supplementary Figures 3-5). Adequacy of imputations for valid statistical inference were verified based on the recommended measures of missing data information of *fraction of missing information* (λ) and *relative increase in variance due to nonresponse* (r)[84, 85] (See Supplemental File 2). Further details of the model specification are in Supplementary file 1. We included participants who had at least 11 pairs of consecutive days of data in the final analytic sample as this provided sufficient amount of data to 1) ensure model convergence and improve reliability and accuracy of the estimates, particularly the random effects and their variances[86-89], and 2) adequately infer participants' habitual exercise level by considering at least three weeks' worth of tracking to compute the weekly exercise frequency. Finally as a post-hoc analysis, we tested the possible influence of type of endometriosis diagnosis by including this categorical variable in the 2 models described above. We conducted the data analyses using R[90] and the glmmTMB package for the GLMMs.[79, 80] Statistical significance level was set at $p<0.05$ for all analyses.

RESULTS

Sample Characteristics. Sample characteristics are provided in Table 1. Participants (N=1,009) had on average 89.6 days of data available for analysis (SD=62.8, Range=22-841,

IQR=31). Participants collectively represent 38 countries, with a wide age range (14-63 years), and varying education and employment status. Almost 70% (N=702) had laparoscopic confirmation of their diagnosis, 19.8% (N=200) had a clinician diagnosis, and 10.6% (N=107) had suspected endometriosis (i.e., "I think I have endometriosis (know the symptoms, no doctor)"). Scores from the VAS, SF-36, and NHS-II Scales are provided in Table 2. The overall prevalence of having used a non-prescription pain medication use was 49.35%, opioid-based medication use was reported by 11.19% of the participants, and similarly use of opioid and paracetamol/acetaminophen combination medications were reported by 11.39% of the participants (See Table 1).

Pain symptom patterns. Mean daily pain score was 4.48 (SD=7.11, 0-79). Mean person-level daily pain score (i.e.,"mean of means") was 4.82 (SD=4.57, Range=0-34). As shown in Figure 1, moderate intensity was the most frequently reported severity across all body areas (Mean=49.3%, SD=22.2), and pelvic pain was the most prevalent area, followed by back pain and gastrointestinal pain (See Figure 1).

Habitual exercise patterns. Mean weekly exercise frequency was 1.43/week (SD=1.54, Range=0-6.87/week, IQR=2.21), 21.3% (N=215) of the sample had an exercise frequency of at least three times per week, and 38.5% (388) of the sample did not engage in any regular exercise (i.e., <1/week). Consequently, 40.2% (N=406) of the sample had a mean exercise frequency of 1-2 times per week. Prevalence of the 10 most frequently reported exercise modalities in the sample are depicted in Figure 2. Walking was the most common modality, reported by 50.94% of the participants, followed by yoga (30.82%), and muscle strength/endurance training activities (24.38%). Yoga and stretching exercises were collectively reported by almost 45% of the sample.

Association of day-level pain to exercise. Tables 3 and 4 display results of the GLMMs estimating day-level total pain score and difference. Coefficients for the model interaction terms indicated a small but statistically significant moderation of previous-day exercise by habitual exercise frequency (RR=0.96 for total pain score and -0.14 for pain score difference, $p<0.05$; See Figure 3). Further inspection of this interaction indicated a mean typical exercise frequency of ~3 times/week as the point after which previous-day exercise began to be associated with favorable pain outcomes (e.g., a decrease from the predicted mean score) on the following day, adjusted for other day-level and person-level factors (Figure 3). This suggests that, participants who typically engage in exercise 3 or more times per week were more likely to report lower pain score and smaller increases (or larger decreases) in pain the day after an exercise bout, compared to not having exercised the previous day. On the other hand, those who exercised less frequently or none were more likely to report higher levels of pain and larger increases (or smaller decreases) in pain 1 day after an exercise bout compared to not having exercised the day before.

Variability in estimated pain scores. There was substantial between-person variability in average day-level pain scores, based on the statistically significant random effect of participant in the models (See Tables 3 and 4, also depicted in Figure 4). The significance of this random effect can further be quantified through a restricted likelihood ratio test (RLRT) based on simulations from the model sample distribution, [91, 92] yielding an observed likelihood ratio (RLRT =7183.3, p-value < 0.0001). These collectively indicate substantial between-individual variability in daily pain experience contributing to the total model pain variance.

Post-hoc analyses. Inclusion of diagnosis type in the model did not have an influence on the results based on the non-significant B coefficients (p=0.48 and p=0.59 for pain score and p=0.70 and p=0.27 for difference in pain score). There were no differences across the 3 groups

with respect to either daily total pain score or difference (χ^2 = 1415.1, df = 1438, p-value = 0.661) (See Supplementary Tables 2 and 3 for full results).

DISCUSSION

Summary of findings. We leveraged 90,382 days of mHealth self-tracking data from 1,009 women with endometriosis to investigate the association between exercise behavior and day-level fluctuations in pain. For the average individual, the association between previous-day exercise to pain was moderated by their habitual exercise frequency, i.e., the frequency with which they engaged in exercised in a typical week. This effect was consistent across participants and independent of person-level covariates. There further was substantial between-person heterogeneity in day-level pain patterns. To our knowledge, this is the first study to quantify the association between day-level pain symptoms and exercise in an international sample of women with endometriosis and to identify habitual weekly exercise frequency as a moderator of this relationship.

Moderation effects. Previous-day exercise was associated with more favorable pain outcomes for participants who engaged in regular exercise at least 3 times per week in our sample. In contrast, those who engaged in regular exercise less than twice a week were more likely to experience pain symptoms on days after having engaged in exercise. This is in line with the national physical activity guidelines [93], which recommend aerobic exercise at least 3 times per week and muscle-strengthening exercise at least twice per week.[94] However, there are no specific recommendations for endometriosis in the current guidelines; and systematic reviews recommend "overall, general exercise" without further details due to lack adequate research on the optimal dose of exercise for endometriosis pain.[4, 36] Our findings provide preliminary

evidence for informing exercise recommendations for endometriosis pain management (i.e., prevention or reduction), specifically for targeting those who are at greater risk for insufficient regular exercise due to acute exacerbations in their pain after exercise. This moderation effect suggests that an individual might need to develop a regular, sustained exercise behavior (i.e., habit) to start experiencing the favorable pain outcomes associated with acute bouts of exercise. Nevertheless, future experimental studies are warranted for a comprehensive investigation of this question.

Patterns of pain symptoms. Our findings of moderate pain in pelvis as the most frequently reported pain are in line with those from others on endometriosis[95] and various chronic pain conditions.[96, 97] The distribution of the total daily pain scores was right-skewed (i.e., extreme scores on the higher ends of the range) with a mean score that was on the lower end of the range. This could partly be due to the data collection method which includes not just days where the participant experienced pain but also days without pain. Indeed, our participants on average did not report or experience any pain 6.25% of the time. In contrast, traditional study designs typically rely on recall of past pain experience aggregated over a period of time (e.g., past week, month) and ask the participant to report their average or highest pain severity over this period.[98, 99] Such recall-based techniques are prone to peak-and-end effects,[100] and catastrophizing or other similar biases.[99, 101] Recruitment from clinical referral points is a common practice and this has been attributed to higher normative scores in the literature,[98] as opposed to more even distributions of pain symptomology among community-based samples.[102] Self-tracking facilitates documentation of not only severe pain, but also mild, moderate, and no pain instances, therefore enabling a more realistic representation of the pain experience as it dynamically unfolds over time. This can reduce the likelihood of over-

representing severe cases, which is a potential limitation attributed to data collected at point of contact in clinical settings.[17] However, it is difficult to make direct comparisons with other studies given the different pain measures, warranting further research.

Patterns of exercise behavior. The mean weekly exercise frequency in the study sample was 1.43/week (SD=1.57, IQR=2.29), with only 24.5 % (N=202) engaging in exercise at least three times a week. This suggests that individuals with endometriosis might be at increased risk for physical inactivity[93, 94], which is a risk factor for various comorbidities [103] and further linked to exacerbation of chronic pain.[104, 105] These collectively underscore the need to focus efforts on promoting regular exercise in women with endometriosis. Notably, yoga and stretching were reported collectively by almost half of the sample within Phendo. This could indicate that participants use these approaches for pain relief, in line with a previous study reporting efficacy of hatha yoga.[33] Nevertheless, participants overall tracked a wide range of exercise modalities across the intensity spectrum (e.g., yoga vs running/cycling) as helpful for their symptoms, suggesting between-individual variability in responses to a given exercise type or intensity. This can be targeted through individualized exercise prescriptions,[25, 106] providing precedence for undertaking a precision approach for pain self-management in endometriosis. Various individualization approaches (e.g., adaptive treatment strategies,[107] micro-randomized trials,[108] just-in-time adaptive interventions [109]) have been investigated for intervening on health behaviors, including PA.[5, 108] It would be opportune to implement a similar N-of-1 intervention approach for identifying person-specific optimal "dose" of exercise based on its parameters to target endometriosis pain symptoms.

Consideration of person-level factors. Another novel finding in our study was the similar point estimates for the effect of exercise on pain outcomes between those with clinician/surgical-

versus suspected diagnosis of endometriosis. Endometriosis is difficult to diagnose, with a 7.6-year delay between symptom onset and its surgical diagnosis.[20, 110, 111] Endometriosis patients further face insurance-related challenges in accessing healthcare for their condition.[15, 112] The participants without a formal diagnosis might have sought medical care for their symptoms but not received the needed care (e.g., diagnostic testing, referral to a specialist), received false negative diagnostic tests results,[110] or lacked adequate access to healthcare. This finding underscores the need for further research in endometriosis that considers self-report of endometriosis symptoms, instead of limiting to patients with a physician referral or relying on secondary data sources (e.g., electronic health records).

Novel methodological contributions. In contrast to other existing questionnaires in the literature, the self-tracking items in Phendo measure momentary and daily pain symptoms and exercise –a time interval for which there are no standard validated, commonly used measures designed for frequent sampling. Computation of a composite pain has been proposed by others[63] as this circumvents numerous limitations in current pain assessment approaches, including lack of a standard single outcome that can be used universally,[61] or a validated instrument that captures all the constructs of persistent pain.[113] There is furthermore a lack of endometriosis-specific pain measures for repeated assessments, thus the heuristic composite pain measure allowed consideration of two dimensions of pain simultaneously in our analyses. The scores in the current study sample were moderately correlated with those from the pelvic-abdominal VAS and the SF-36 bodily pain measure, which were also similarly correlated with each other ($r=0.46$, $p<0.0001$). Nevertheless, future directions include evaluation of this measure in larger samples for its reliability and validity via a nomological network-based analysis.

Limitations. We acknowledge several limitations of this study, including reliance on self-

reports for the type of endometriosis diagnosis and exercise behavior.First, we used a binary measure of exercise in our analyses and did not have sufficient details on duration or intensity for inclusion in the analyses as potential moderators. Similarly, we did not have granular daily data on pain medication use, as such it was not investigated as a potential covariate in the analyses. In addition to medications, future studies could consider other pain management approaches for comparison to exercise, given previous research suggesting endometriosis patients report using a variety of symptom management techniques.[45] Next, our sample consisted primarily of White, non-Hispanic women who are relatively consistent mHealth technology users and furthermore can understand English to use the App. Therefore the results might differ among other groups including non-English speakers or those without an interest in mHealth use for self-management or monitoring.

Conclusion

In this study, we provide evidence that habitual exercise frequency is a potential moderator of the association between pain symptoms and previous-day exercise in endometriosis, indicating that those who regularly exercise at least ~3 times per week are less likely to report pain symptoms after having exercised on the previous day. Individuals with endometriosis are significantly more likely to have higher all-cause healthcare utilization and direct health care costs than those without endometriosis, including twice the prevalence of opioid prescriptions for pain management [23] and prolonged duration of prescriptions.[22] While guidelines recommend prescribing exercise for management of pain in clinical populations, endometriosis (or general chronic) pain-specific recommendations to guide patients and providers on measurable parameters (time, type, intensity, and frequency) are lacking. This

warrants future studies investigating the effects of both acute and chronic exercise on endometriosis pain with a focus on various types, intensities and durations.

Author Contributions

IE conceptualized the study, conducted the data analyses, and prepared the first draft of the manuscript. SLG and ENH were responsible for data acquisition, curation and management. NE acquired the funding and provided the mHealth infrastructure for the study (Phendo App). NE and SB provided guidance on the study design and data analyses. SB, NE, SLG and ENH critically reviewed and provided feedback on the manuscript.

Funding

Funding for the work is provided by a postdoctoral fellowship from the Data Science Institute at Columbia University and an award from the National Library of Medicine (R01 LM013043). We are grateful to the Phendo participants.

Competing Interests

All authors report no conflicts of interest.

Data availability statement

Data are available on reasonable request.

1. Available at https://itunes.apple.com/us/app/phendo/id1145512423
2. Available at https://play.google.com/store/apps/details?id=com.appliedinformaticsinc.phendo

References

1. Ambrose KR, Golightly YM. Physical exercise as non-pharmacological treatment of chronic pain: Why and when. Best Pract Res Clin Rheumatol. 2015;29(1):120-30. Epub 2015/05/23. doi: 10.1016/j.berh.2015.04.022. PubMed PMID: 26267006.

2. Rice D, Nijs J, Kosek E, Wideman T, Hasenbring MI, Koltyn K, et al. Exercise-Induced Hypoalgesia in Pain-Free and Chronic Pain Populations: State of the Art and Future Directions. The Journal of Pain. 2019;20(11):1249-66. doi: https://doi.org/10.1016/j.jpain.2019.03.005.

3. Lemmens J, De Pauw J, Van Soom T, Michiels S, Versijpt J, van Breda E, et al. The effect of aerobic exercise on the number of migraine days, duration and pain intensity in migraine: a systematic literature review and meta-analysis. J Headache Pain. 2019;20(1):16. Epub 2019/02/16. doi: 10.1186/s10194-019-0961-8. PubMed PMID: 30764753; PubMed Central PMCID: PMCPMC6734345.

4. Armour M, Ee CC, Naidoo D, Ayati Z, Chalmers KJ, Steel KA, et al. Exercise for dysmenorrhoea. Cochrane Database of Systematic Reviews. 2019;(9). doi: 10.1002/14651858.CD004142.pub4. PubMed PMID: CD004142.

5. Rabbi M, Aung MS, Gay G, Reid MC, Choudhury T. Feasibility and acceptability of mobile phone–based auto-personalized physical activity recommendations for chronic pain self-management: pilot study on adults. Journal of medical Internet research. 2018;20(10):e10147.

6. Sevel L, Boissoneault J, Alappattu M, Bishop M, Robinson M. Training endogenous pain modulation: a preliminary investigation of neural adaptation following repeated exposure to clinically-relevant pain. Brain Imaging and Behavior. 2020;14(3):881-96. doi: 10.1007/s11682-018-0033-8.

7. Gordon R, Bloxham S, editors. A systematic review of the effects of exercise and physical activity on non-specific chronic low back pain. Healthcare; 2016: Multidisciplinary Digital Publishing Institute.

8. Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH. Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. Cochrane Database of Systematic Reviews. 2017;(4).

9. Zhang R, Chomistek AK, Dimitrakoff JD, Giovannucci EL, Willett WC, Rosner BA, et al. Physical activity and chronic prostatitis/chronic pelvic pain syndrome. Med Sci Sports Exerc. 2015;47(4):757-64. doi: 10.1249/MSS.0000000000000472. PubMed PMID: 25116086.

10. Pinto A, Di Raimondo D, Tuttolomondo A, Buttà C, Milio G, Licata G. Effects of physical exercise on inflammatory markers of atherosclerosis. Current pharmaceutical design. 2012;18(28):4326-49.

11. Garatachea N, Molinero O, Martínez-García R, Jimenez-Jimenez R, Gonzalez-Gallego J, Marquez S. Feelings of well being in elderly people: relationship to physical activity and physical function. Archives of Gerontology and Geriatrics. 2009;48(3):306-12.

12. Tennfjord MK, Gabrielsen R, Tellum T. Effect of physical activity and exercise on endometriosis-associated symptoms: a systematic review. BMC Women's Health. 2021;21(1):355. doi: 10.1186/s12905-021-01500-4.

13. Evans S, Fernandez S, Olive L, Payne LA, Mikocka-Walus A. Psychological and mind-body interventions for endometriosis: A systematic review. J Psychosom Res. 2019;124:109756. Epub 2019/08/25. doi: 10.1016/j.jpsychores.2019.109756. PubMed PMID: 31443810.

14. Mira TAA, Buen MM, Borges MG, Yela DA, Benetti-Pinto CL. Systematic review and meta-analysis of complementary treatments for women with symptomatic endometriosis. Int J Gynaecol Obstet. 2018;143(1):2-9. Epub 2018/06/27. doi: 10.1002/ijgo.12576. PubMed PMID: 29944729.

15. Fourquet J, Gao X, Zavala D, Orengo JC, Abac S, Ruiz A, et al. Patients' report on how endometriosis affects health, work, and daily life. Fertil Steril. 2010;93(7):2424-8. doi: 10.1016/j.fertnstert.2009.09.017. PubMed PMID: 19926084.

16. Schliep KC, Mumford SL, Peterson CM, Chen Z, Johnstone EB, Sharp HT, et al. Pain typology and incident endometriosis. Hum Reprod. 2015;30(10):2427-38. Epub 2015/08/11. doi: 10.1093/humrep/dev147. PubMed PMID: 26269529.

17. De Graaff A, D'hooghe T, Dunselman G, Dirksen C, Hummelshoj L, Consortium WE, et al. The significant effect of endometriosis on physical, mental and social wellbeing: results from an international cross-sectional survey. Human reproduction. 2013;28(10):2677-85.

18. Simoens S, Dunselman G, Dirksen C, Hummelshoj L, Bokor A, Brandes I, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. Human Reproduction. 2012;27(5):1292-9.

19. Soliman AM, Coyne KS, Gries KS, Castelli-Haley J, Snabes MC, Surrey ES. The effect of endometriosis symptoms on absenteeism and presenteeism in the workplace and at home. Journal of managed care & specialty pharmacy. 2017;23(7):745-54.

20. Simoens S, Dunselman G, Dirksen C, Hummelshoj L, Bokor A, Brandes I, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. Hum Reprod. 2012;27(5):1292-9. Epub 2012/03/17. doi: 10.1093/humrep/des073. PubMed PMID: 22422778.

21. The Practice Committee of the American Society for Reproductive Medicine. Treatment of pelvic pain associated with endometriosis: a committee opinion. 2014 2014/04/01/. Report No.: 0015-0282 Contract No.: 4.

22. Lamvu G, Soliman AM, Manthena SR, Gordon K, Knight J, Taylor HS. Patterns of prescription opioid use in women with endometriosis: evaluating prolonged use, daily dose, and concomitant use with benzodiazepines. Obstetrics and gynecology. 2019;133(6):1120.

23. Soliman AM, Surrey ES, Bonafede M, Nelson JK, Vora JB, Agarwal SK. Health care utilization and costs associated with endometriosis among women with medicaid insurance. Journal of managed care & specialty pharmacy. 2019;25(5):566-72.

24. Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain—United States, 2016. Jama. 2016;315(15):1624-45.

25. Sluka KA, Frey-Law L, Hoeger Bement M. Exercise-induced pain and analgesia? Underlying mechanisms and clinical translation. Pain. 2018;159 Suppl 1(Suppl 1):S91-S7. doi: 10.1097/j.pain.0000000000001235. PubMed PMID: 30113953.

26. Tour J, Löfgren M, Mannerkorpi K, Gerdle B, Larsson A, Palstam A, et al. Gene-to-gene interactions regulate endogenous pain modulation in fibromyalgia patients and healthy controls-antagonistic effects between opioid and serotonin-related genes. Pain. 2017;158(7):1194-203. Epub 2017/03/11. doi: 10.1097/j.pain.0000000000000896. PubMed PMID: 28282362; PubMed Central PMCID: PMCPMC5472004.

27. Bobinski F, Teixeira JM, Sluka KA, Santos ARS. Interleukin-4 mediates the analgesia produced by low-intensity exercise in mice with neuropathic pain. Pain. 2018;159(3):437-50. Epub 2017/11/16. doi: 10.1097/j.pain.0000000000001109. PubMed PMID: 29140923; PubMed Central PMCID: PMCPMC5812806.

28. Montenegro ML, Bonocher CM, Meola J, Portella RL, Ribeiro-Silva A, Brunaldi MO, et al. Effect of Physical Exercise on Endometriosis Experimentally Induced in Rats. Reproductive sciences (Thousand Oaks, Calif). 2018:1933719118799205. Epub 2018/09/21. doi: 10.1177/1933719118799205. PubMed PMID: 30231769.

29. Stratton P, Berkley KJ. Chronic pelvic pain and endometriosis: translational evidence of the relationship and implications. Human reproduction update. 2011;17(3):327-46. Epub 2010/11/23. doi: 10.1093/humupd/dmq050. PubMed PMID: 21106492.

30. Park S-J, Yong M-S, Na S-S. Effect of exercise on the expression of nerve growth factor in the spinal cord of rats with induced osteoarthritis. Journal of physical therapy science. 2015;27(8):2551-4. Epub 2015/08/21. doi: 10.1589/jpts.27.2551. PubMed PMID: 26357438.

31. Karasawa Y, Yamada K, Iseki M, Yamaguchi M, Murakami Y, Tamagawa T, et al. Association between change in self-efficacy and reduction in disability among patients with chronic pain. PLOS ONE. 2019;14(4):e0215404. doi: 10.1371/journal.pone.0215404.

32. Armour M, Sinclair J, Chalmers KJ, Smith CA. Self-management strategies amongst Australian women with endometriosis: a national online survey. BMC Complementary and Alternative Medicine. 2019;19(1):17. doi: 10.1186/s12906-019-2431-x.

33. Gonçalves AV, Barros NF, Bahamondes L. The Practice of Hatha Yoga for the Treatment of Pain Associated with Endometriosis. Journal of Alternative & Complementary Medicine. 2017;23(1):45-52. doi: 10.1089/acm.2015.0343. PubMed PMID: 120746246.

34. Ricci E, Viganò P, Cipriani S, Chiaffarino F, Bianchi S, Rebonato G, et al. Physical activity and endometriosis risk in women with infertility or pain: Systematic review and meta-analysis. Medicine. 2016;95(40):e4957-e. doi: 10.1097/MD.0000000000004957. PubMed PMID: 27749551.

35. Carpenter SE, Tjaden B, Rock JA, Kimball A. The effect of regular exercise on women receiving danazol for treatment of endometriosis. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics. 1995;49(3):299-304. Epub 1995/06/01. PubMed PMID: 9764869.

36. Bonocher CM, Montenegro ML, Rosa ESJC, Ferriani RA, Meola J. Endometriosis and physical exercises: a systematic review. Reproductive biology and endocrinology : RB&E. 2014;12:4. Epub 2014/01/08. doi: 10.1186/1477-7827-12-4. PubMed PMID: 24393293; PubMed Central PMCID: PMCPMC3895811.

37. Naugle KM, Fillingim RB, Riley JL, 3rd. A meta-analytic review of the hypoalgesic effects of exercise. The journal of pain : official journal of the American Pain Society. 2012;13(12):1139-50. Epub 2012/11/08. doi: 10.1016/j.jpain.2012.09.006. PubMed PMID: 23141188.

38. Janal MN, Colt EWD, Clark WC, Glusman M. Pain sensitivity, mood and plasma endocrine levels in man following long-distance running: Effects of naloxone. Pain. 1984;19(1):13-25. doi: https://doi.org/10.1016/0304-3959(84)90061-7.

39. Droste C, Greenlee MW, Schreck M, Roskamm H. Experimental pain thresholds and plasma beta-endorphin levels during exercise. Medicine & Science in Sports & Exercise. 1991;23(3):334-42. doi: 10.1249/00005768-199103000-00012.

40. Scheef L, Jankowski J, Daamen M, Weyer G, Klingenberg M, Renner J, et al. An fMRI study on the acute effects of exercise on pain processing in trained athletes. PAIN. 2012;153(8).

41. Hoffman MD, Hoffman DR. Exercisers Achieve Greater Acute Exercise-Induced Mood Enhancement Than Nonexercisers. Archives of Physical Medicine and Rehabilitation. 2008;89(2):358-63. doi: 10.1016/j.apmr.2007.09.026.

42. Hallgren M, Moss ND, Gastin P. Regular exercise participation mediates the affective response to acute bouts of vigorous exercise. J Sports Sci Med. 2010;9(4):629-37. Epub 2010/01/01. PubMed PMID: 24149790; PubMed Central PMCID: PMCPMC3761821.

43. Chen Y-C, Chen C, Martínez RM, Etnier JL, Cheng Y. Habitual physical activity mediates the acute exercise-induced modulation of anxiety-related amygdala functional connectivity. Scientific Reports. 2019;9(1):19787. doi: 10.1038/s41598-019-56226-z.

44. May M, Junghaenel DU, Ono M, Stone AA, Schneider S. Ecological Momentary Assessment Methodology in Chronic Pain Research: A Systematic Review. J Pain. 2018;19(7):699-716. Epub 2018/01/31. doi: 10.1016/j.jpain.2018.01.006. PubMed PMID: 29371113.

45. Ensari I, Pichon A, Lipsky-Gorman S, Bakken S, Elhadad N. Augmenting the Clinical Data Sources for Enigmatic Diseases: A Cross-Sectional Study of Self-Tracking Data and Clinical Documentation in Endometriosis. Applied Clinical Informatics. 2020;11(05):769-84.

46. Urteaga I, McKillop M, Elhadad N. Learning endometriosis phenotypes from patient-generated data. npj Digital Medicine. 2020;3(1):88. doi: 10.1038/s41746-020-0292-9.

47. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. npj Digital Medicine. 2020;3(1):21. doi: 10.1038/s41746-020-0224-8.

48. McKillop M, Voigt N, Schnall R, Elhadad N. Exploring self-tracking as a participatory research activity among women with endometriosis. Journal of Participatory Medicine. 2016.

49. McKillop M, Mamykina L, Elhadad N, editors. Designing in the Dark: Eliciting Self-tracking Dimensions for Understanding Enigmatic Disease. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; 2018: ACM.

50. Frost MH, Reeve BB, Liepa AM, Stauffer JW, Hays RD. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? Value Health. 2007;10 Suppl 2:S94-s105. Epub 2007/11/13. doi: 10.1111/j.1524-4733.2007.00272.x. PubMed PMID: 17995479.

51. Anthoine E, Moret L, Regnault A, Sébille V, Hardouin J-B. Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. Health Qual Life Outcomes. 2014;12:176-. doi: 10.1186/s12955-014-0176-2. PubMed PMID: 25492701.

52. US Department of Health Human Services. Guidance for industry-Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

53. Lomas J, Pickard L, Mohide A. Patient versus clinician item generation for quality-of-life measures: the case of language-disabled adults. Medical Care. 1987:764-9.

54. Melzack R. The McGill Pain Questionnaire: Major properties and scoring methods. PAIN. 1975;1(3).

55. Jones KR, Vojir CP, Hutt E, Fink R. Determining mild, moderate, and severe pain equivalency across pain-intensity tools in nursing home residents. J Rehabil Res Dev. 2007;44(2):305-14. Epub 2007/06/07. doi: 10.1682/jrrd.2006.05.0051. PubMed PMID: 17551881.

56. Bestel E, Gotteland J-P, Donnez J, Taylor RN, Garner EI. Linzagolix for Endometriosis-Associated Pain: Lipid Changes After 52 Weeks of Treatment [25B]. Obstetrics & Gynecology. 2020;135:25S. doi: 10.1097/01.AOG.0000663180.46470.c9. PubMed PMID: 00006250-202005001-00082.

57. Serlin RC, Mendoza TR, Nakamura Y, Edwards KR, Cleeland CS. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. Pain. 1995;61(2):277-84. Epub 1995/05/01. doi: 10.1016/0304-3959(94)00178-h. PubMed PMID: 7659438.

58. Adams P, Murnane EL, Elfenbein M, Wethington E, Gay G. Supporting the Self-Management of Chronic Pain Conditions with Tailored Momentary Self-Assessments. Proc SIGCHI Conf Hum Factor Comput Syst. 2017;2017:1065-77. doi: 10.1145/3025453.3025832. PubMed PMID: 30310887.

59. Lee RR, Rashid A, Ghio D, Thomson W, Cordingley L. "Seeing Pain Differently": A Qualitative Investigation Into the Differences and Similarities of Pain and Rheumatology Specialists' Interpretation of Multidimensional Mobile Health Pain Data From Children and Young People With Juvenile Idiopathic Arthritis. JMIR Mhealth Uhealth. 2019;7(7):e12952. Epub 2019/07/04. doi: 10.2196/12952. PubMed PMID: 31267979; PubMed Central PMCID: PMCPMC6632104.

60. Jamison RN, Raymond SA, Levine JG, Slawsby EA, Nedeljkovic SS, Katz NP. Electronic diaries for monitoring chronic pain: 1-year validation study. Pain. 2001;91(3):277-85.

61. Bouhassira D, Attal N. All in one: Is it possible to assess all dimensions of any pain with a simple questionnaire? PAIN. 2009;144(1).

62. Boonstra AM, Stewart RE, Köke AJA, Oosterwijk RFA, Swaan JL, Schreurs KMG, et al. Cut-Off Points for Mild, Moderate, and Severe Pain on the Numeric Rating Scale for Pain in Patients with Chronic Musculoskeletal Pain: Variability and Influence of Sex and Catastrophizing. Front Psychol. 2016;7:1466-. doi: 10.3389/fpsyg.2016.01466. PubMed PMID: 27746750.

63. Pilitsis JG, Fahey M, Custozzo A, Chakravarthy K, Capobianco R. Composite score is a better reflection of patient response to chronic pain therapy compared with pain intensity alone. Neuromodulation: Technology at the Neural Interface. 2021;24(1):68-75.

64. Dirks JF, Wunder J, Kinsman R, McElhinny J, Jones NF. A Pain Rating Scale and a Pain Behavior Checklist for Clinical Use: Development, Norms, and the Consistency Score. Psychotherapy and Psychosomatics. 1993;59(1):41-9.

65. Faurholt-Jepsen M, Munkholm K, Frost M, Bardram JE, Kessing LV. Electronic self-monitoring of mood using IT platforms in adult patients with bipolar disorder: A systematic review of the validity and evidence. BMC Psychiatry. 2016;16(1):7. doi: 10.1186/s12888-016-0713-0.

66. Charter RA. Sample size requirements for precise estimates of reliability, generalizability, and validity coefficients. J Clin Exp Neuropsychol. 1999;21(4):559-66. Epub 1999/11/07. doi: 10.1076/jcen.21.4.559.889. PubMed PMID: 10550813.

67. Knell G, Gabriel KP, Businelle MS, Shuval K, Wetter DW, Kendzor DE. Ecological Momentary Assessment of Physical Activity: Validation Study. J Med Internet Res. 2017;19(7):e253. doi: 10.2196/jmir.7602.

68. Swendeman D, Comulada WS, Koussa M, Worthman CM, Estrin D, Rotheram-Borus MJ, et al. Longitudinal Validity and Reliability of Brief Smartphone Self-Monitoring of Diet, Stress, and Physical Activity in a Diverse Sample of Mothers. JMIR Mhealth Uhealth. 2018;6(9):e176. Epub 2018/09/27. doi: 10.2196/mhealth.9378. PubMed PMID: 30249576; PubMed Central PMCID: PMCPMC6231816.

69. Katapally TR, Chu LM. Digital epidemiological and citizen science methodology to capture prospective physical activity in free-living conditions: a SMART Platform study. BMJ Open. 2020;10(6):e036787. Epub 2020/07/01. doi: 10.1136/bmjopen-2020-036787. PubMed PMID: 32595163; PubMed Central PMCID: PMCPMC7322321.

70. Zink J, Belcher BR, Dzubur E, Ke W, O'Connor S, Huh J, et al. Association Between Self-Reported and Objective Activity Levels by Demographic Factors: Ecological Momentary Assessment Study in Children. JMIR Mhealth Uhealth. 2018;6(6):e150. doi: 10.2196/mhealth.9592.

71. Ensari I, Horan E, Bakken S, Elhadad N. Evaluation of a disease-specific mHealth exercise measure for daily self-tracking MedRXiv. 2022;Under review.

72. Kuntz JL, Young DR, Saelens BE, Frank LD, Meenan RT, Dickerson JF, et al. Validity of the Exercise Vital Sign Tool to Assess Physical Activity. Am J Prev Med. 2021;60(6):866-72. Epub 2021/03/31. doi: 10.1016/j.amepre.2021.01.012. PubMed PMID: 33781618; PubMed Central PMCID: PMCPMC8154650.

73. Wolf AM, Hunter DJ, Colditz GA, Manson JE, Stampfer MJ, Corsano KA, et al. Reproducibility and validity of a self-administered physical activity questionnaire. Int J Epidemiol. 1994;23(5):991-9. Epub 1994/10/01. doi: 10.1093/ije/23.5.991. PubMed PMID: 7860180.

74. Vitonis AF, Vincent K, Rahmioglu N, Fassbender A, Buck Louis GM, Hummelshoj L, et al. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonization Project: II. Clinical and covariate phenotype data collection in endometriosis research. Fertility and sterility. 2014;102(5):1223-32. Epub 2014/09/22. doi: 10.1016/j.fertnstert.2014.07.1244. PubMed PMID: 25256930.

75. Jones G, Kennedy S, Barnard A, Wong J, Jenkinson C. Development of an endometriosis quality-of-life instrument: The Endometriosis Health Profile-30. Obstet Gynecol. 2001;98(2):258-64. Epub 2001/08/17. doi: 10.1016/s0029-7844(01)01433-8. PubMed PMID: 11506842.

76. McHorney CA, Ware JE, Lu JFR, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36): III. Tests of Data Quality, Scaling Assumptions, and Reliability across Diverse Patient Groups. Medical Care. 1994;32(1):40-66.

77. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc. 2011;43(8):1575-81.

78. Hu M-C, Pavlicova M, Nunes EV. Zero-inflated and hurdle models of count data with extra zeros: examples from an HIV-risk reduction intervention trial. Am J Drug Alcohol Abuse. 2011;37(5):367-75. doi: 10.3109/00952990.2011.597280. PubMed PMID: 21854279.

79. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R journal. 2017;9(2):378-400.

80. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. Modeling zero-inflated count data with glmmTMB. bioRxiv. 2017:132753. doi: 10.1101/132753.

81. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software; Vol 1, Issue 3 (2011). 2011.

82. Van Buuren S. Flexible imputation of missing data: CRC press; 2018.

83. Bondarenko I, Raghunathan T. Graphical and numerical diagnostic tools to assess suitability of multiple imputations and imputation models. Statistics in Medicine. 2016;35(17):3007-20. doi: https://doi.org/10.1002/sim.6926.

84. Rubin DB. The Calculation of Posterior Distributions by Data Augmentation: Comment: A Noniterative Sampling/Importance Resampling Alternative to the Data Augmentation Algorithm for Creating a Few Imputations When Fractions of Missing Information Are Modest: The SIR Algorithm. Journal of the American Statistical Association. 1987;82(398):543-6. doi: 10.2307/2289460.

85. Rubin DB. Multiple imputation for nonresponse in surveys: John Wiley & Sons; 2004.

86. Schunck R. Cluster Size and Aggregated Level 2 Variables in Multilevel Models. A Cautionary Note. 2016. 2016;10(1). Epub 2016-07-20. doi: 10.12758/mda.2016.005.

87. Bell B, Ferron J, Kromrey J, editors. Cluster Size in Multilevel Models: The Impact of Sparse Data Structures on Point and Interval Estimates in Two-Level Models2008.

88. Austin PC, Leckie G. The effect of number of clusters and cluster size on statistical power and Type I error rates when testing random effects variance components in multilevel linear and logistic regression models. Journal of Statistical Computation and Simulation. 2018;88(16):3151-63. doi: 10.1080/00949655.2018.1504945.

89. Snijders TAB. Power and sample size in multilevel modeling. In: Everitt B, Howell D, editors. Encyclopedia of Statistics in Behavioral Science. 3: Wiley; 2006. p. 1570–3.

90. Core Team R. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria1997.

91. Scheipl F, Greven S, Küchenhoff H. Size and power of tests for a zero random effect variance or polynomial regression in additive and linear mixed models. Computational statistics & data analysis. 2008;52(7):3283-99.

92. Crainiceanu CM, Ruppert D. Likelihood ratio tests in linear mixed models with one variance component. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2004;66(1):165-85.

93. US Department of Health Human Services Physical activity guidelines advisory committee scientific report. 2018.

94. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. Jama. 2018;320(19):2020-8.

95. Warzecha D, Szymusik I, Wielgos M, Pietrzak B. The Impact of Endometriosis on the Quality of Life and the Incidence of Depression-A Cohort Study. Int J Environ Res Public Health. 2020;17(10). Epub 2020/05/28. doi: 10.3390/ijerph17103641. PubMed PMID: 32455821; PubMed Central PMCID: PMCPMC7277332.

96. Becker N, Thomsen AB, Olsen AK, Sjogren P, Bech P, Eriksen J. Pain epidemiology and health related quality of life in chronic non-malignant pain patients referred to a Danish multidisciplinary pain center. Pain. 1997;73(3):393-400. doi: 10.1016/s0304-3959(97)00126-7. PubMed PMID: WOS:000071429200014.

97. Bouhassira D, Lantéri-Minet M, Attal N, Laurent B, Touboul C. Prevalence of chronic pain with neuropathic characteristics in the general population. Pain. 2008;136(3):380-7.

98. Nicholas MK, Asghari A, Blyth FM. What do the numbers mean? Normative data in chronic pain measures. Pain. 2008;134(1):158-73. doi: https://doi.org/10.1016/j.pain.2007.04.007.

99. Dansie EJ, Turk DC. Assessment of patients with chronic pain. Br J Anaesth. 2013;111(1):19-25. doi: 10.1093/bja/aet124. PubMed PMID: 23794641.

100. Schneider S, Stone AA, Schwartz JE, Broderick JE. Peak and end effects in patients' daily recall of pain and fatigue: a within-subjects analysis. J Pain. 2011;12(2):228-35. doi: 10.1016/j.jpain.2010.07.001. PubMed PMID: 20817615.

101. De Boer M, Struys M, Versteegen G. Pain-related catastrophizing in pain patients and people with pain in the general population. European journal of pain. 2012;16(7):1044-52.

102. Ehde DM, Gibbons LE, Chwastiak L, Bombardier CH, Sullivan MD, Kraft GH. Chronic pain in a large community sample of persons with multiple sclerosis. Multiple Sclerosis Journal. 2003;9(6):605-11. doi: 10.1191/1352458503ms939oa.

103. Katzmarzyk PT, Powell KE, Jakicic JM, Troiano RP, Piercy K, Tennant B, et al. Sedentary Behavior and Health: Update from the 2018 Physical Activity Guidelines Advisory Committee. Med Sci Sports Exerc. 2019;51(6):1227-41. doi: 10.1249/MSS.0000000000001935. PubMed PMID: 31095080.

104. Basen-Engquist K, Scruggs S, Jhingran A, Bodurka DC, Lu K, Ramondetta L, et al. Physical activity and obesity in endometrial cancer survivors: associations with pain, fatigue, and physical functioning. Am J Obstet Gynecol. 2009;200(3):288.e1-.e2888. Epub 2008/12/25. doi: 10.1016/j.ajog.2008.10.010. PubMed PMID: 19110220.

105. Dansie EJ, Turk DC, Martin KR, Van Domelen DR, Patel KV. Association of Chronic Widespread Pain With Objectively Measured Physical Activity in Adults: Findings From the National Health and Nutrition Examination Survey. The Journal of Pain. 2014;15(5):507-15. doi: https://doi.org/10.1016/j.jpain.2014.01.489.

106. Polaski AM, Phelps AL, Kostek MC, Szucs KA, Kolber BJ. Exercise-induced hypoalgesia: A meta-analysis of exercise dosing for the treatment of chronic pain. PloS one. 2019;14(1):e0210418-e. doi: 10.1371/journal.pone.0210418. PubMed PMID: 30625201.

107. Almirall D, Compton SN, Gunlicks-Stoessel M, Duan N, Murphy SA. Designing a pilot sequential multiple assignment randomized trial for developing an adaptive treatment strategy. Statistics in medicine. 2012;31(17):1887-902. Epub 2012/03/23. doi: 10.1002/sim.4512. PubMed PMID: 22438190; PubMed Central PMCID: PMCPMC3399974.

108. Klasnja P, Smith S, Seewald NJ, Lee A, Hall K, Luers B, et al. Efficacy of Contextually Tailored Suggestions for Physical Activity: A Micro-randomized Optimization Trial of HeartSteps. Annals of behavioral medicine : a publication of the Society of Behavioral Medicine. 2018. Epub 2018/09/08. doi: 10.1093/abm/kay067. PubMed PMID: 30192907.

109. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. Annals of behavioral medicine : a publication of the Society of Behavioral Medicine. 2018;52(6):446-62. Epub 2016/09/25. doi: 10.1007/s12160-016-9830-8. PubMed PMID: 27663578; PubMed Central PMCID: PMCPMC5364076.

110. Falcone T, Mascha E. The elusive diagnostic test for endometriosis. Fertility and sterility. 2003;80(4):886-8.

111. Marian S, Hermanowicz-Szamatowicz K. Endometriosis–a decade later–still an enigmatic disease. What is the new in the diagnosis and treatment? Gynecological Endocrinology. 2020;36(2):104-8.

112. Fourquet J, Zavala DE, Missmer S, Bracero N, Romaguera J, Flores I. Disparities in healthcare services in women with endometriosis with public vs private health insurance. American Journal of Obstetrics & Gynecology. 2019;221(6):623.e1-.e11. doi: 10.1016/j.ajog.2019.06.020.

113. Grimmer-Somers K, Vipond N, Kumar S, Hall G. A review and critique of assessment instruments for patients with persistent pain. Journal of pain research. 2009;2:21.

Table 1. Study Sample Characteristics.

Characteristic (N)	Mean (SD) / Frequency (%)
Age (827)	31.0 (7.26), Median=30.6 (MAD=7.41), Range= 14.3-62.9
BMI (787)	25.9 (6.98), Median=24.1 (MAD=4.74), Range= 16.01-72.24
Type of endometriosis diagnosis	
Surgery (702)	69.57 %
Clinician (200)	19.82 %
Self-diagnosis (107)	10.60 %
Work Environment	
Home (218)	26.42 %
Outside (570)	69.09 %
Unknown (221)	21.29 %
Living environment	
Rural (129)	15.27 %
Suburban (340)	41.21 %
Urban (363)	44.00 %
Unknown (161)	19.5 %
Relationship status	
Married/domestic partnership (442)	53.57 %
Separated/divorced (28)	3.39 %
Single/never married (310)	37.57 %

Unknown (229)	22.69 %
Education Level	
College or higher (547)	66.30 %
High school graduate or less (74)	8.96 %
Some college (209)	25.33 %
Unknown (179)	17.7 %
Employment Status	
Employed (541)	65.57 %
Not employed (120)	14.54 %
Student (129)	15.63 %
Unknown (219)	21.70 %
Race/Ethnicity	
White, Non-Hispanic (699)	84.72 %
Black, Non-Hispanic (20)	2.42 %
Asian (22)	2.6 %
Native American (6)	0.72 %
Hispanic (38)	4.6 %
Other (51)	6.18 %
Unknown (173)	17.14 %
Country of Residence	
United States (444)	44.0 %
United Kingdom (83)	8.22 %
Canada (75)	7.43 %

Australia (59)	5.84 %
Germany (38)	3.76 %
New Zealand (34)	3.36 %
Other (69)	6.83 %
Unknown (207)	20.51 %

Table 2. Sample Study Scores on Standard Measures of Pain and Exercise.

EPQ-S Measures (N)	**Mean (SD)**
SF-36 Bodily Pain (375)	35.47 (22.33)
Pelvic-abdominal pain VAS (316)	7.37 (1.97)
NHS-II PA Scale Total Weekly Minutes (359)	175.2 (280.2)
NHS-II PA Scale Total Weekly EE (359)	16.13 (30.37)

Table 3. Results of the regression model estimating day-level total pain score (N=1,009).

Conditional Random Effects		**Variance (95% CI)**	
Participant		1.09 (0.98, 1.21)	
Conditional Fixed Effects	**Rate Ratio (95% CI)**	**Log Odds (SE)**	***z*-score**
Intercept	4.26*** (3.26, 5.56)	1.45*** (0.13)	10.82
Menstrual Status	1.29*** (1.25, 1.32)	0.25*** (0.01)	20.31
Previous Day Pain	1.02*** (1.02, 1.03)	0.02*** (0.00)	29.69
Body Mass Index	1.01* (1.00, 1.02)	0.01 (0.00)	2.02
Mean Weekly Exercise Frequency	0.93* (0.89, 0.97)	-0.06** (0.02)	-2.96
Previous Day exercise	1.10* (1.05, 1.15)	0.09**(0.15)	3.88
Some College Education Level	0.87 (0.83, 1.56)	0.13 (0.15)	0.86
College or Higher Education Level	0.93 (0.66, 1.16)	-0.13 (0.14)	-0.92
Mean Weekly Exercise Frequency * Previous Day exercise	0.96** (0.95, 0.98)	-0.03** (0.01)	-3.37
Zero Inflation Terms	**Rate Ratio (95% CI)**	**Log Odds (SE)**	***z*-score**
Intercept	0.17 (0.16, 0.18)	-1.73***(0.02)	-62.96
Same Day Exercise	5.34 (5.01, 5.68)	1.67*** (0.03)	52.53

95% CI=95% Confidence Interval.*p<0.05 ** p <0.001, ***p<0.0001. Previous day pain and BMI were sample mean-centered. BMI and education level were kept as covariates in the model based on their significant associations with mean day-level pain scores (Pearson's r=0.15 for BMI and Kruskal-Wallis χ^2 = 18.061 for education level, p<0.001).

Table 4. Results of the regression model estimating pain score difference (N=1,009).

Conditional Random Effects	**Variance (95% CI)**
Participant (Intercept)	9.16 (8.28, 10.13)
Residual	26.83

Conditional Fixed Effects	**B coefficient (SE)**	**95% CI**	***z*-score**
Intercept	2.70*** (0.51)	1.68, 3.72	5.29
Menstrual Status	1.47*** (0.09)	1.28, 1.66	15.43
Previous Day Pain	-0.86*** (0.01)	-0.87, -0.85	-143.43
Body Mass Index	0.05* (0.01)	0.01, 0.10	2.86
Mean Weekly Exercise Frequency	-0.27** (0.08)	-0.44, -0.10	-3.12
Previous Day Exercise	0.92** (0.18)	0.56, 1.27	5.08
Some College Education Level	-0.84 (0.62)	-2.11, 0.42	-1.35
College or Higher Education Level	-2.07** (0.52)	-3.10, -1.03	-3.96
Mean Weekly Exercise Frequency * Previous Day Exercise	-0.14* (0.06)	-0.26, -0.01	-2.22

Zero Inflation Terms	**B coefficient**	**95% CI**	***z*-score**
Intercept	-0.91*** (0.01)	-0.93, -0.88	-63.84
Same Day Exercise	0.70*** (0.02)	0.66, 0.75	32.09

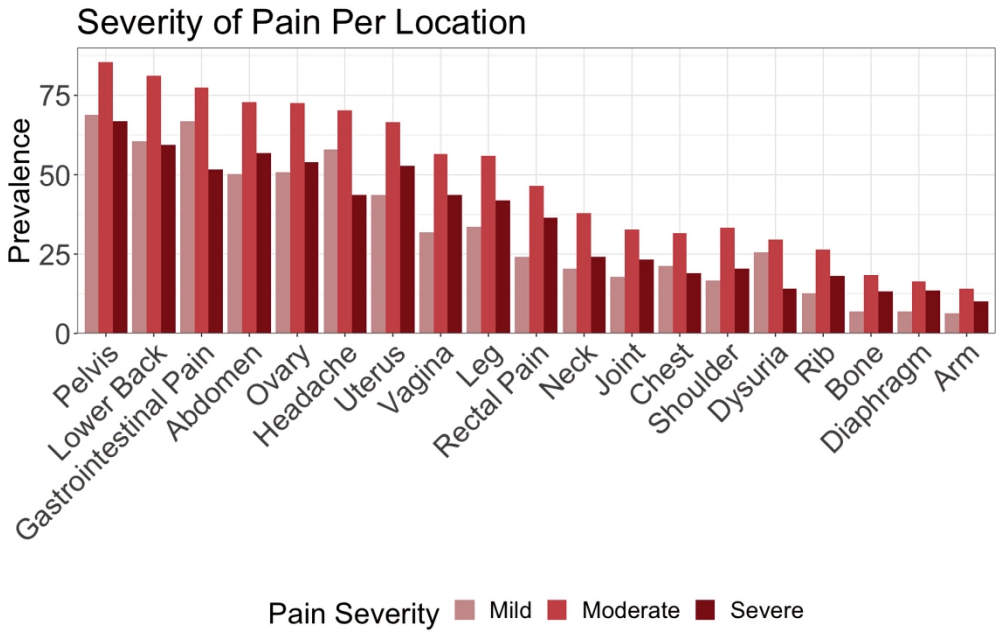
SE= Standard Error. *p<0.05 ** p <0.001, ***p<0.0001. Previous day pain and BMI were sample mean-centered. BMI and education level were kept as covariates in the model based on their significant associations with mean day-level pain scores (Pearson's r=0.15 for BMI and Kruskal-Wallis χ^2 = 18.061 for education level, p<0.001).

Figure 1. Prevalence of pain severity by location reported among participants (i.e., unique counts of body area-severity per participant). Moderate intensity was the most frequently tracked across all body areas (14.1%-85.4%).

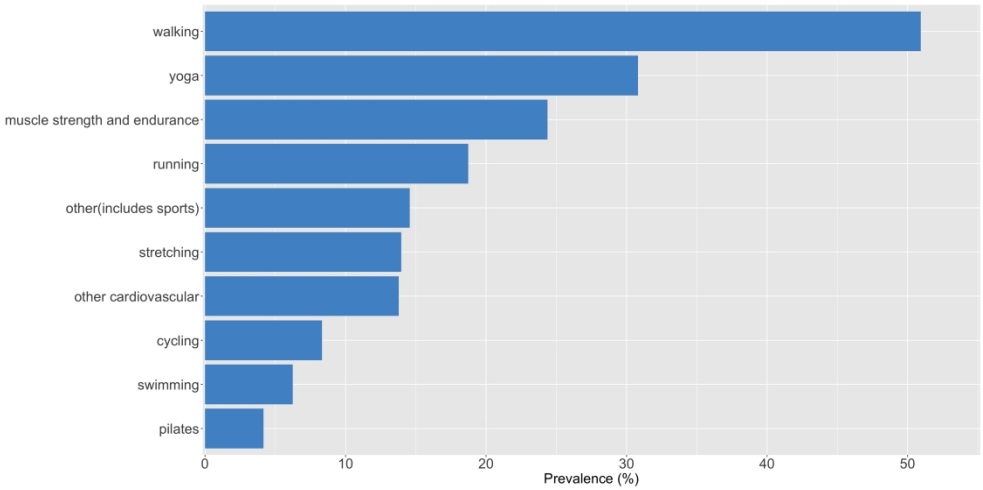
Figure 2. Prevalence of self-reported exercise modalities in the study sample. "Other cardiovascular" category include activities such as dancing, aerobics and using the elliptical machine. "Muscle strength and endurance" category includes activities such as weight lifting and calisthenics. "Other exercise" category includes sports activities such as skiing and soccer, multi-modal exercises (e.g., high intensity interval training of both cardiovascular and muscular endurance), or those that did not fit into the other categories (e.g., stabilizing or balancing exercises, wii fit or other home based fitness activities).

Panel Figure 3. Moderation of effect of previous-day exercise by habitual exercise levels (X axes). Y axes represent predicted day-level total scores (top) and differences (bottom) in pain. Shaded areas depict 95% confidence intervals. At approximately 3 times/week of regular exercise, previous day exercise starts to be associated with more favorable pain outcomes on the following day (i.e., decrease from the model predicted mean scores), adjusted for other day-level and person level factors.

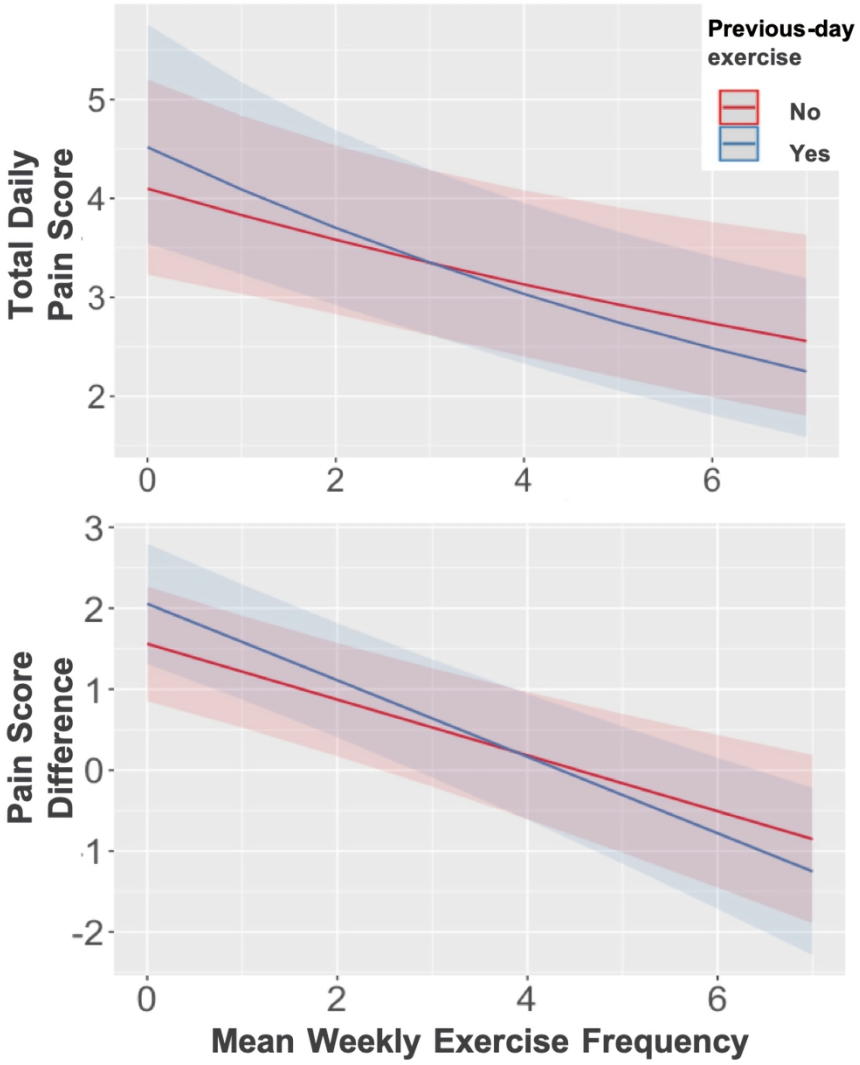
Figure 4. Plot of the random effect of the participant on total day pain scores estimated from the multilevel model (N=1,009). Y-axis represents the range of estimated average pain scores for each participant. Each black dot represents one participant's mean (i.e., random intercept), grey lines indicate 95% confidence intervals. Distribution of points across the x-axis indicate large variability across individuals (i.e., between-group variance), and the grey lines indicate the within-person variability in daily scores over time.



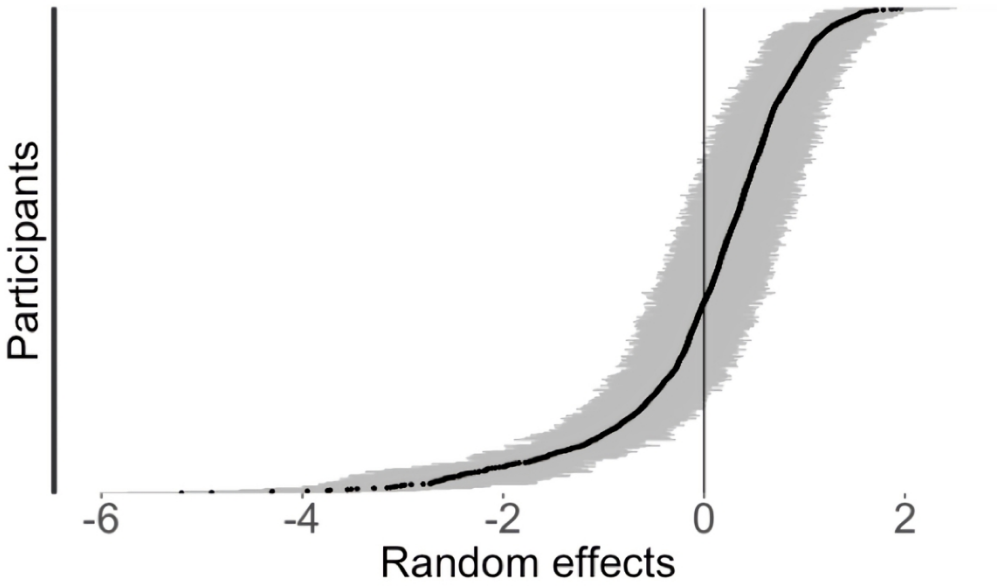
776x493mm (72 x 72 DPI)



1411x705mm (72 x 72 DPI)



694x750mm (72 x 72 DPI)



571x339mm (72 x 72 DPI)

Supplementary File 1. Missing Data Imputations.

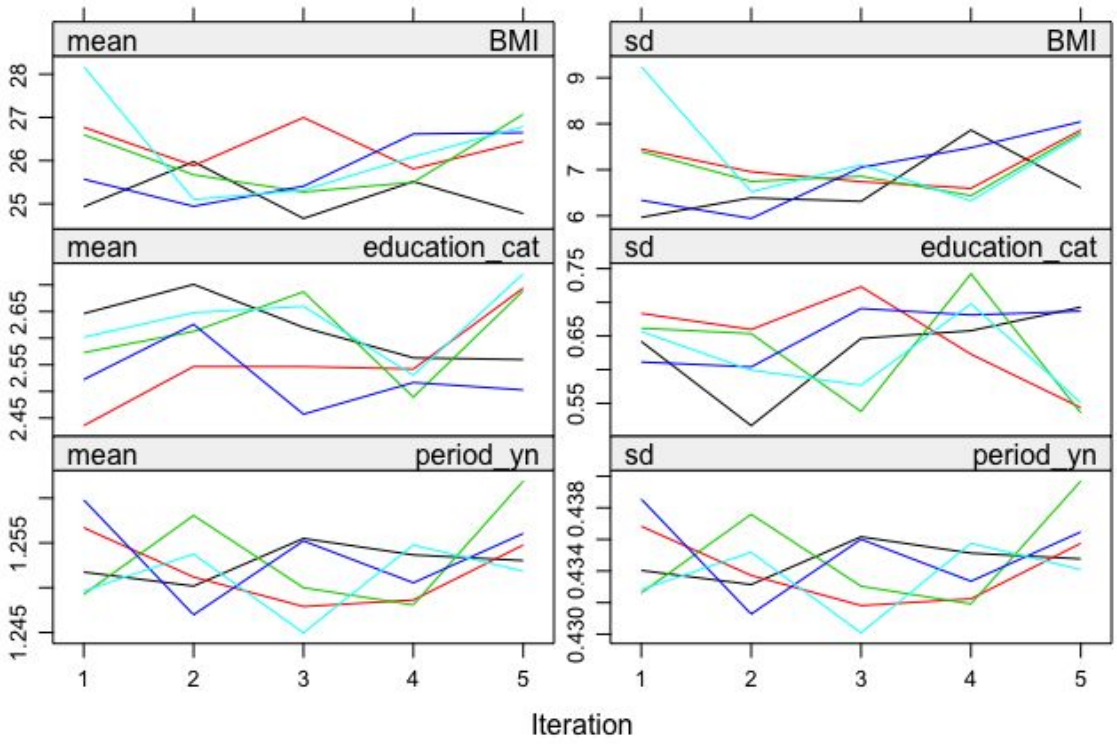
Phendo is an observational research app and participants do not receive prompts from the research team to track any given item at a certain time. They are free to track (or not track) any given item as they wish. Consequently, missingness in the data occurs due to a variety of possible reasons that are not always known or easy to distinguish. For example, a period not tracked for a day could mean that the participant did not have a period, or they chose not to track, or did not use the app at all that day. To circumvent this issue, we took several measures. First, we limited data to days for which the participant tracked their pain, exercise and menstrual status at least once, as a proxy for app use. Next, we assigned a score of zero for pain on days where the participant had tracked exercise or menstrual status but not pain. This approach is motivated by 2 reasons. First, the nature of the pain question in Phendo (i.e., "Where is the pain?", "How severe is the pain?") assumes the participants to track when they feel pain and therefore a "No Pain" response is neither available in the app nor would make sense. Second, multiple imputation methods impute such that the resulting imputations are limited to the observed values and distributions. Thus by default it would omit the possibility of a zero in the resultant pain score distribution, which increases risk of overestimation of the scores in the sample.

BMI (calculated from participant reported height and weight) and education level were missing for 22% and 19% of the participants, respectively, and menstrual status was missing (i.e., not tracked) 22% of the time in the dataset. We imputed these 3 variables using multivariate imputations by chained equations [1] according to the heteroscedastic linear two-level structure of the data (i.e., hierarchical where, participant is the clustering variable) following standard multilevel multiple imputation methods. [1-4] We used two-level predictive mean matching for

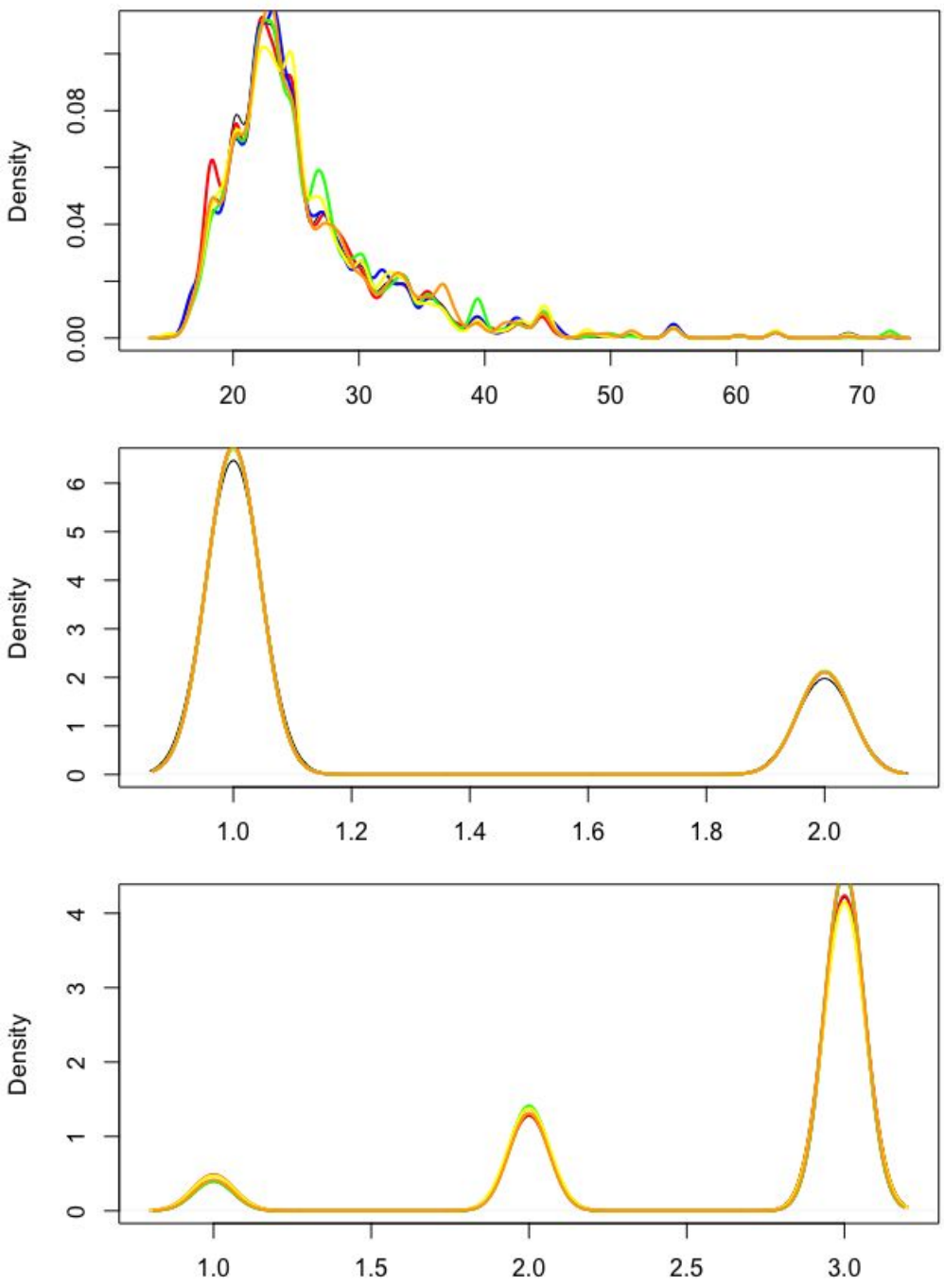
BMI and education level, which is a semi-parametric imputation method that limits imputations to the observed values and can preserve non-linear relations in the observed data, therefore the imputations do not deviate from the observed distribution[5] and two-level logistic regression for imputing menstrual status, using the rest of the dataset as the predictors. As per published recommendations,[1, 2] we also included the raw pain variable (i.e., with the missing values) as a predictor, to account for the possibility of an association between the missingness pattern of pain to these imputed variables. To assess the plausibility of the imputations and any significant deviance from the structure of the raw, non-imputed data, we inspected the imputation convergence plots, distributions of the imputed variables which are provided in Supplementary Figures 3 and 4.

Model specification. We used a zero-inflated negative binomial (ZINB) distribution when modeling the total pain outcome, as it has been demonstrated to provide the best fit for outcomes with over-dispersion and zero-inflation.[6-8] ZINB models consider two sources of zero observations: "sampling zeros" that are part of the underlying sampling distribution (i.e., negative binomial) and "structural zeros" that cannot score anything other than zero (i.e., participant did not track).[6] This virtue of the ZINB models allows for specification of the imputed zeros and prevents the risk of over-estimating effects and generates more conservative estimates for predictors of interest by estimating a separate zero-inflation term, as well as conditional model.[6] We specified the zero-inflation term such that it was dependent on the exercise variable for the day, in addition to specifying an overall general zero-inflation structure in the outcome through inclusion of an intercept, based on recommendations. [8] Menstrual status was not a significant predictor of zero-inflation and therefore removed from the zero-inflation term during the modeling process. We included participants who had at least 11 pairs of

consecutive days of data in the final analytic sample as this provided sufficient amount of data to 1) ensure model convergence and improve reliability and accuracy of the estimates, particularly the random effects and their variances[9-12], and 2) adequately infer participants' habitual weekly exercise frequency by considering at least three weeks' worth of tracking to compute the weekly exercise frequency.



Supplementary Figure 3. Convergence plots for the 3 imputed variables (BMI, top; education, middle; menstrual status, bottom) with means on the left and standard deviations on the right side of the panel. Plots indicate healthy convergence based on lack significant trend and the streams mingling well right from the start throughout the 5 iterations (x-axis).



Supplementary Figure 4. Density plots of the marginal distributions of BMI (top), menstrual status (middle), and education category (bottom) of raw, non-imputed data and 5 iterations of the imputed data. Close super-imposition of the curves indicate that the imputed data distributions match those of raw data.

Supplementary File 2. Imputation Model diagnostics.

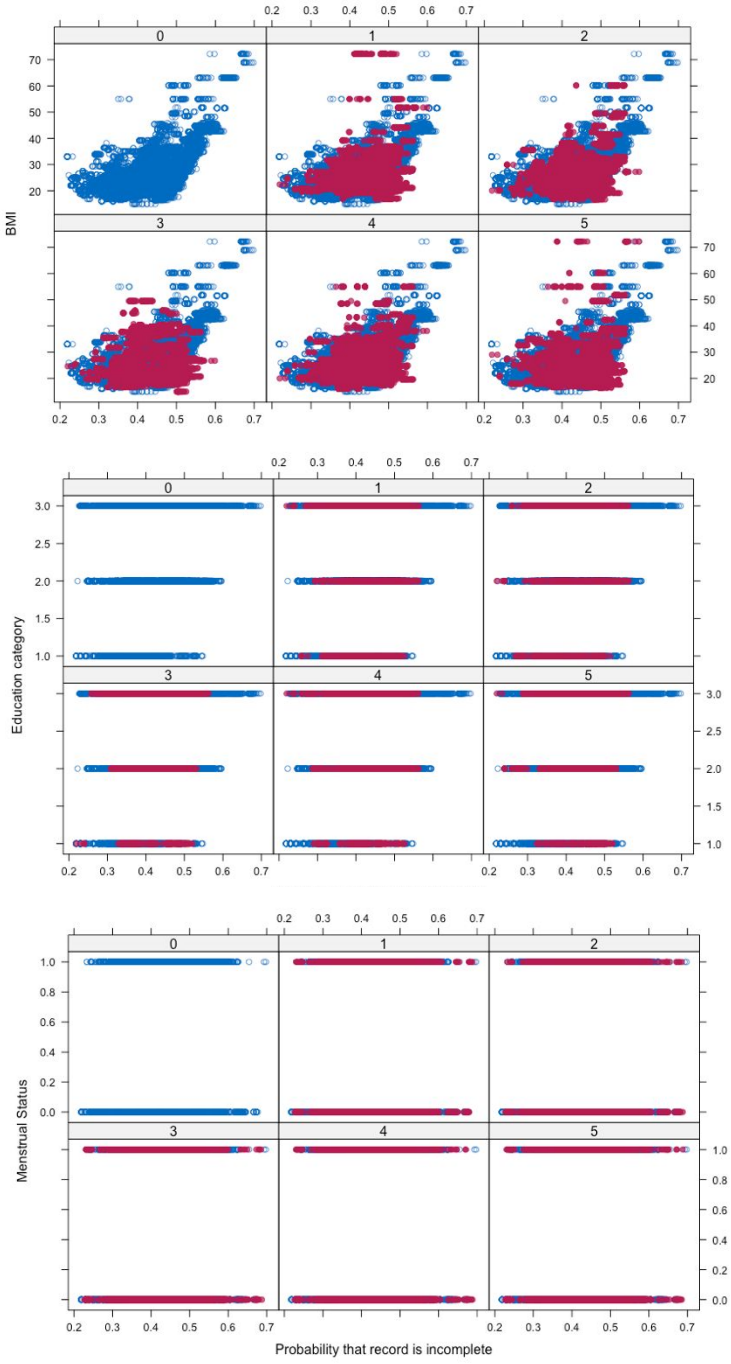
Appropriateness and plausibility of the estimates from imputed models were inspected following published guidelines. First, we used measures of missing data information to assess pooled estimate variances. The fraction of missing information (λ) is interpreted as the proportion of variation in the parameter of interest due to the missing data. The relative increase in variance due to nonresponse (r) is interpreted as the proportional increase in the sampling variance of the parameter of interest that is due to the missing data. Values of λ over 0.5 indicate that the influence of the imputation model on the results is larger than that of the complete-data model, suggesting potential problems in the imputations. Supplementary Table 1 provides results of these variance estimates, indicating satisfactory imputation and model fit.

Supplementary Table 1. Measures of Missing data information

	Total Pain Score		**Difference in Pain**	
Conditional Fixed Effects	**λ**	***r***	**λ**	***r***
Intercept	0.21	0.27	0.23	0.31
Menstrual Status	0.13	0.15	0.19	0.23
Previous Day Pain	0.01	0.01	0.00	0.00
Body Mass Index	0.13	0.15	0.23	0.31
Mean Weekly Exercise Frequency	0.00	0.00	0.01	0.01
Previous Day exercise	0.01	0.01	0.00	0.00
Some College Education Level	0.26	0.36	0.35	0.55
College or Higher Education Level	0.23	0.31	0.21	0.28
Mean Weekly Exercise Frequency * Previous Day exercise	0.00	0.00	0.00	0.00
Zero Inflation Terms				
Intercept	0.00	0.00	0.00	0.00
Same Day Exercise	0.00	0.00	0.00	0.00

Next, we inspected propensity scores, which is a more recent and increasingly accepted method for inspecting the suitability of data imputation.[2, 13, 14] The goal is to compare the distributions of observed and imputed data conditional on the missingness probability. Under the missing at random (MAR) assumption, the conditional distributions of the observed and missing data should be similar if the assumed model for creating multiple imputations has a good fit. To do this, we first estimate the probability of each record being incomplete (i.e., "response propensity") in the presence of missing data by conditioning on the response indicators as well as the observed covariates. The probabilities are then averaged over the imputed datasets to obtain stability. Supplementary Figure 3 plots BMI, education category and menstrual status against the propensity score in each dataset. The distributions of the blue

and red points are match up well without significant discrepancies (e.g., mismatch in patterns, imputed data systematically shifted toward one side of the axis).



Supplementary Figure 5. BMI (top), education category (middle) and menstrual status (bottom) plotted against the propensity score in each dataset (0=observed, 1-5=imputed). The distributions of the blue (observed) and red (imputed) points are follow similar patterns.

Supplementary Table 2. Post-hoc analyses with endometriosis diagnosis included as a covariate. Conditional model results of the negative binomial model estimation of day-level total pain score (N=608).

Random Effects	**Variance (95% CI)**
Participant (Intercept)	1.10 (0.99, 1.22)

Fixed Effects	**Log Odds (SE)**	**z-score**
Intercept	1.37*** (0.12)	10.97
Menstrual Status	0.25*** (0.01)	21.40
Previous day Pain	0.02*** (0.01)	21.40
Body Mass Index	0.01* (0.004)	2.81
Mean weekly Exercise Frequency	-0.06** (0.02)	-3.01
Previous day exercise	0.09** (0.02)	3.85
Clinician diagnosis of endometriosis	-0.07 (0.10)	0.01
Self-diagnosis of endometriosis	-0.11 (0.11)	-1.01
Some college education level	0.22 (0.13)	-1.63
College or higher education level	-0.01 (0.12)	-0.12
Mean weekly Exercise Frequency*Previous day exercise	-0.03*** (0.01)	-3.42

SE=Standard Error. *p=0.001, ** p <0.001, ***p<0.0001. B coefficients are rate ratios. BMI =Body Mass Index. BMI and previous day pain were group mean centered.

Supplementary Table 3. Post-hoc analyses with endometriosis diagnosis included as a covariate. Conditional model results of the regression model estimation of pain score difference (N=1009).

Conditional Random Effects	**Variance (95% CI)**
Participant (Intercept)	13.34 (12.09, 14.93)

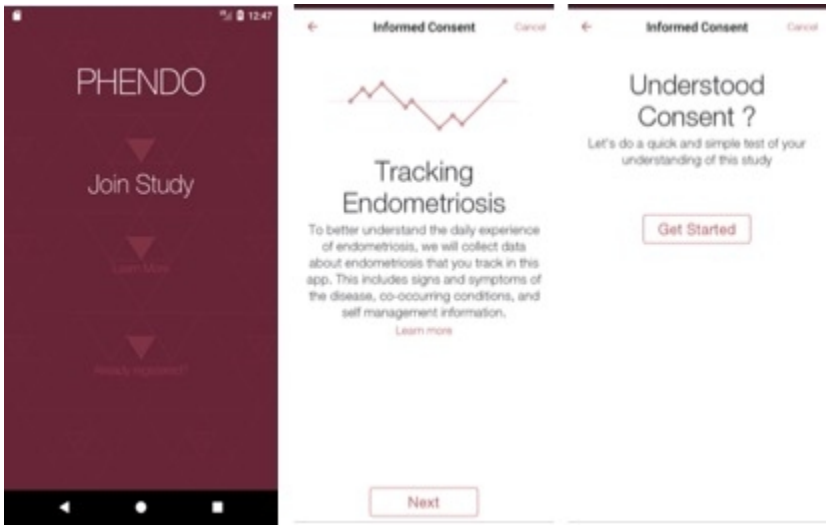
Fixed Effects	**B coefficient (SE)**	**z-score**
Intercept	2.45*** (0.46)	5.22
Menstrual status	1.46*** (0.08)	16.98
Previous day pain	-0.86*** (0.01)	-144.11
Body mass index	0.07* (0.01)	4.47
Mean weekly exercise frequency	-0.27** (0.09)	-3.03
Previous day exercise	0.92*** (0.18)	5.13
Clinician diagnosis of endometriosis	-0.05 (0.32)	-0.16
Self-diagnosis of endometriosis	-0.45 (0.43)	-1.29
Some college education level	-0.30 (0.51)	-0.58
College or higher education level	-1.72** (0.47)	-3.67
Mean weekly exercise frequency*Previous day exercise	-0.14* (0.06)	-2.31

SE=Standard Error. *p<0.05, ** p <0.01, ***p<0.0001. Body Mass Index and previous day pain were group mean centered.

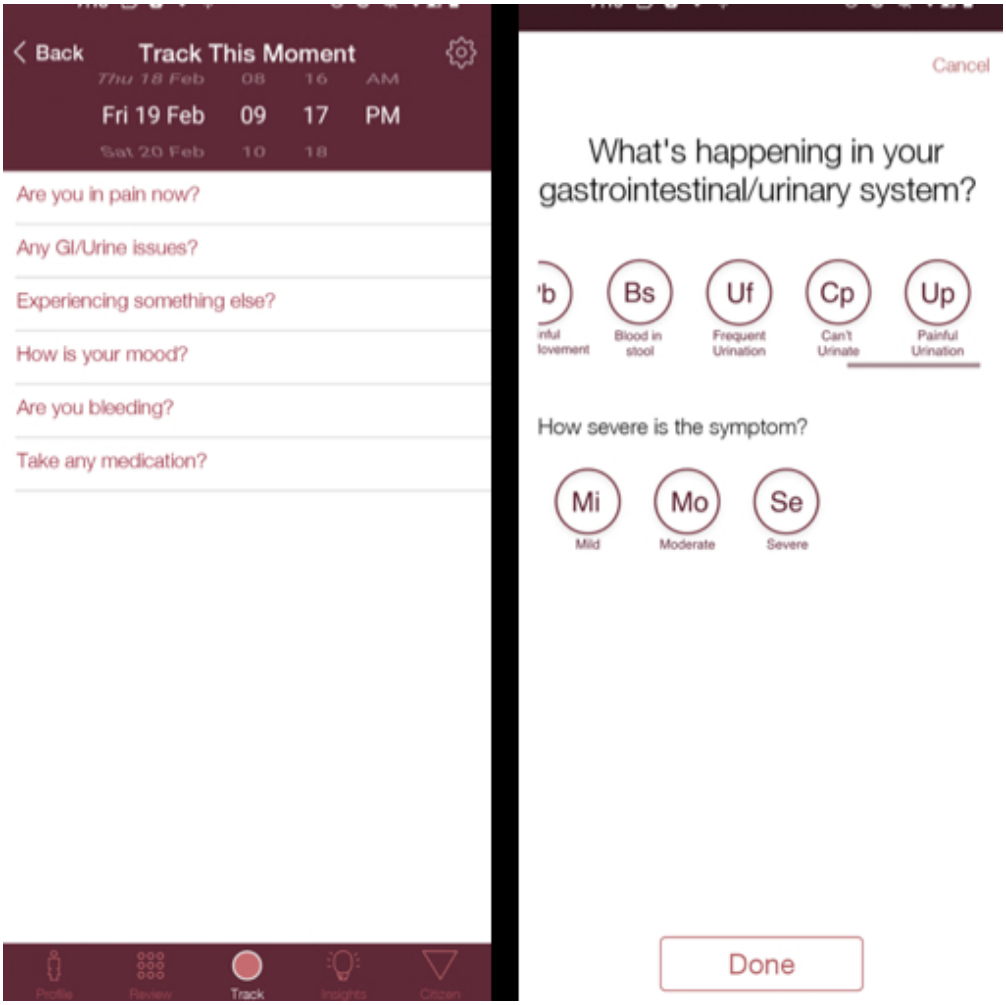
References

1. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software; Vol 1, Issue 3 (2011). 2011.
2. Van Buuren S. Flexible imputation of missing data: CRC press; 2018.
3. Grund S, Lüdtke O, Robitzsch A. Multiple imputation of missing data for multilevel models: Simulations and recommendations. Organizational Research Methods. 2018;21(1):111-49.
4. Barnard J, Rubin DB. Small-Sample Degrees of Freedom with Multiple Imputation. Biometrika. 1999;86(4):948-55.
5. Little RJ. Missing-data adjustments in large surveys. Journal of Business & Economic Statistics. 1988;6(3):287-96.

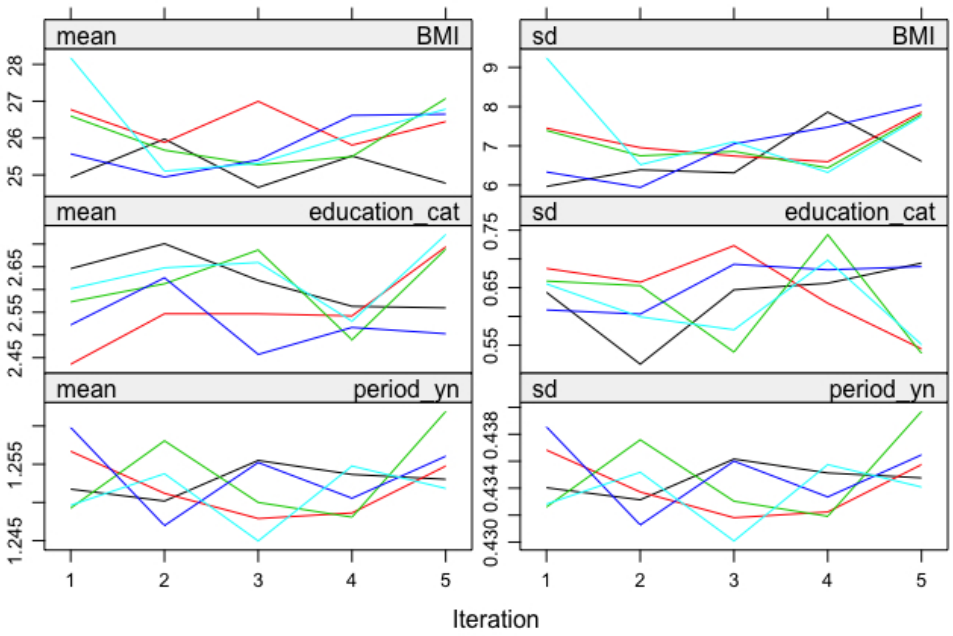
6. Hu M-C, Pavlicova M, Nunes EV. Zero-inflated and hurdle models of count data with extra zeros: examples from an HIV-risk reduction intervention trial. Am J Drug Alcohol Abuse. 2011;37(5):367-75. doi: 10.3109/00952990.2011.597280. PubMed PMID: 21854279.
7. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R journal. 2017;9(2):378-400.
8. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. Modeling zero-inflated count data with glmmTMB. bioRxiv. 2017:132753. doi: 10.1101/132753.
9. Schunck R. Cluster Size and Aggregated Level 2 Variables in Multilevel Models. A Cautionary Note. 2016. 2016;10(1). Epub 2016-07-20. doi: 10.12758/mda.2016.005.
10. Bell B, Ferron J, Kromrey J, editors. Cluster Size in Multilevel Models: The Impact of Sparse Data Structures on Point and Interval Estimates in Two-Level Models2008.
11. Austin PC, Leckie G. The effect of number of clusters and cluster size on statistical power and Type I error rates when testing random effects variance components in multilevel linear and logistic regression models. Journal of Statistical Computation and Simulation. 2018;88(16):3151-63. doi: 10.1080/00949655.2018.1504945.
12. Snijders TAB. Power and sample size in multilevel modeling. In: Everitt B, Howell D, editors. Encyclopedia of Statistics in Behavioral Science. 3: Wiley; 2006. p. 1570–3.
13. Nguyen CD, Carlin JB, Lee KJ. Model checking in multiple imputation: an overview and case study. Emerging Themes in Epidemiology. 2017;14(1):8. doi: 10.1186/s12982-017-0062-6.
14. Bondarenko I, Raghunathan T. Graphical and numerical diagnostic tools to assess suitability of multiple imputations and imputation models. Statistics in Medicine. 2016;35(17):3007-20. doi: https://doi.org/10.1002/sim.6926.



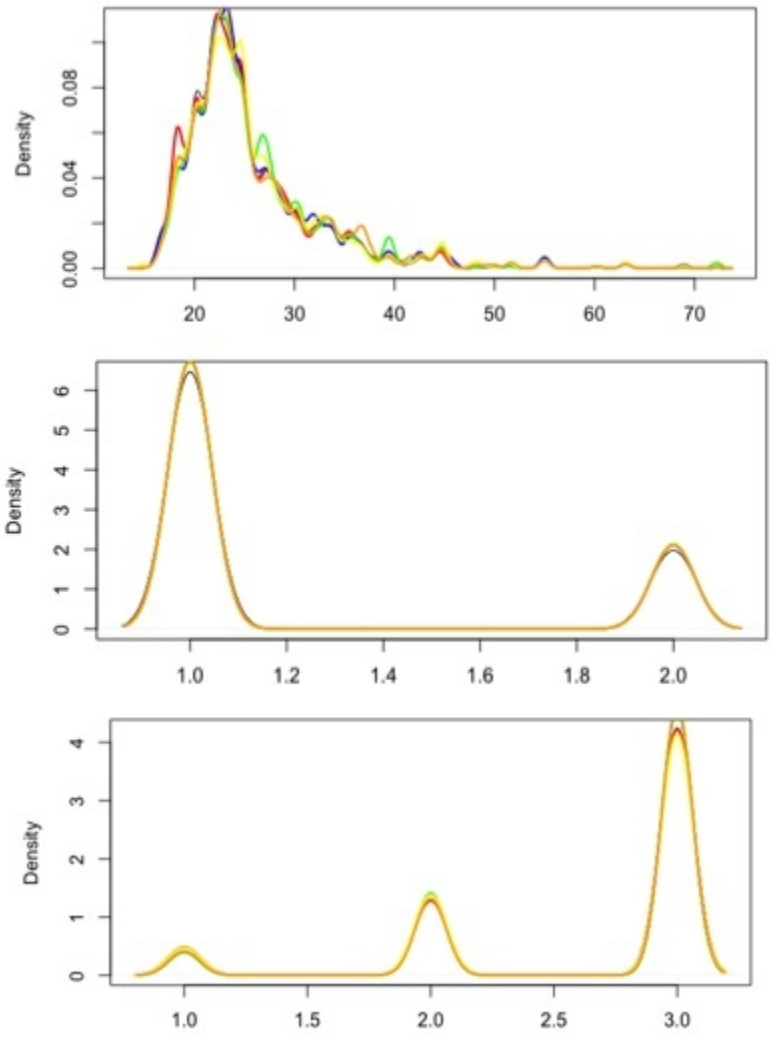
34x22mm (300 x 300 DPI)



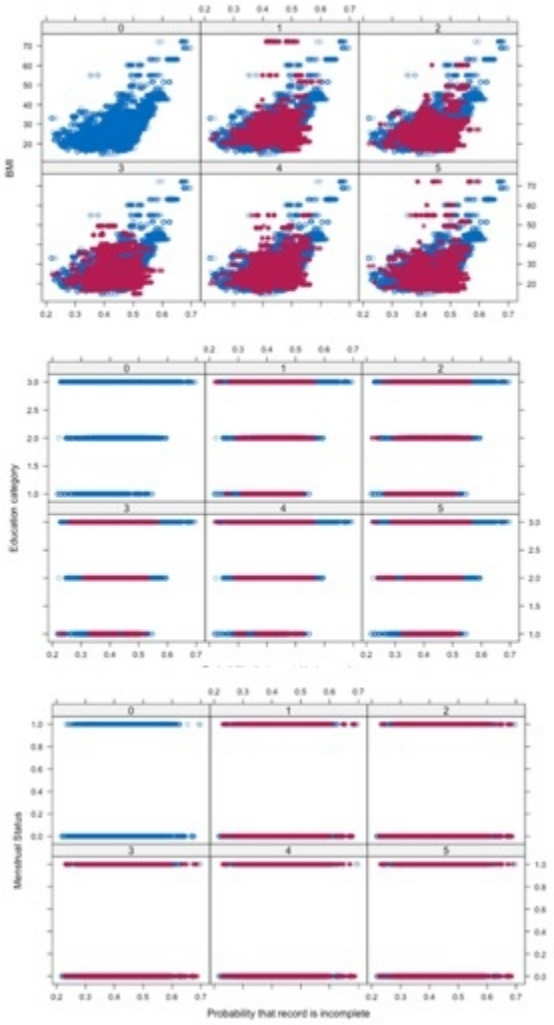
109x109mm (118 x 118 DPI)



142x99mm (118 x 118 DPI)



34x45mm (300 x 300 DPI)



23x44mm (300 x 300 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page # where this item is located:
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1,2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5-6
Objectives	3	State specific objectives, including any prespecified hypotheses	7
Methods			
Study design	4	Present key elements of study design early in the paper	8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	8-9
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	7-8
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed	N/A

		Case-control study—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	9-11
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	9-11
Bias	9	Describe any efforts to address potential sources of bias	10-11
Study size	10	Explain how the study size was arrived at	7-8

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	11-12
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	12-13
		(*b*) Describe any methods used to examine subgroups and interactions	12-13
		(*c*) Explain how missing data were addressed	13, Supplemental Files 1-2
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	12-13
		(*e*) Describe any sensitivity analyses	12, and Supplemental Tables
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	13-14
		(b) Give reasons for non-participation at each stage	13-14
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	13-14, and Table 1
		(b) Indicate number of participants with missing data for each variable of interest	Table 1
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	N/A
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	*N/A*
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	*N/A*
		Cross-sectional study—Report numbers of outcome events or summary measures	14-15
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	15-16, Tables 2 and 3

	(*b*) Report category boundaries when continuous variables were categorized	N/A
	(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	15-16, and Supplemental Tables
Discussion			
Key results	18	Summarise key results with reference to study objectives	16
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	19-20
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	17-19
Generalisability	21	Discuss the generalisability (external validity) of the study results	19-20
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	21

BMJ Open

A cross-sectional mHealth-based investigation of the associations between physical exercise patterns and pain symptoms in individuals with endometriosis

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-059280.R2
Article Type:	Original research
Date Submitted by the Author:	07-Jun-2022
Complete List of Authors:	Ensari, Ipek; Columbia University, Data Science Institute Lipsky-Gorman, Sharon; Columbia University, Department of Biomedical Informatics Horan, Emma; Columbia University, Department of Biomedical Informatics Bakken, Suzanne; Columbia University, School of Nursing Elhadad, Noemie; Columbia University, Department of Biomedical Informatics
Primary Subject Heading:	Health informatics
Secondary Subject Heading:	Public health
Keywords:	PAIN MANAGEMENT, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS, PREVENTIVE MEDICINE, EPIDEMIOLOGY, COMPLEMENTARY MEDICINE

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

A cross-sectional mHealth-based investigation of the associations between physical exercise patterns and pain symptoms in individuals with endometriosis

Running title: Daily exercise and pain patterns in endometriosis

Ipek Ensari[1], PhD, Sharon Lipsky-Gorman[2], MA, Emma Horan[2], BS, Suzanne R. Bakken[3], PhD, FAAN, Noémie Elhadad[2], PhD

[1]Data Science Institute, Columbia University, [2]Department of Biomedical Informatics, Columbia University Irving Medical Center, [3]Columbia University School of Nursing, New York, NY

Author of Correspondence:
Ipek Ensari
Columbia University
Data Science Institute
475 Riverside Dr, Room 320,
New York, NY 10115
Email:ie2145@columbia.edu

Word count: 4,490

Abstract

Objectives: This study investigates the association of daily physical exercise with pain symptoms in endometriosis. We also examined whether an individual's typical weekly (i.e., habitual) exercise frequency influences (i.e., moderates) the relationship between their pain symptoms on a given day (day *t*) and previous-day (day *t-1*) exercise. **Participants:** The sample included 90,382 days of data from 1,009 participants (~85% non-Hispanic white) living with endometriosis across 38 countries. **Study Design:** This was an observational, retrospective study conducted using data from a research mobile app (Phendo) designed for collecting self-reported data on symptoms and self-management of endometriosis. **Primary Outcome Measures:** The two primary outcomes were the composite day-level pain score that includes pain intensity and location, and the change in this score from previous day (Δ-score). We applied generalized linear mixed-level models to examine the effect of previous-day exercise and habitual exercise frequency on these outcomes. We included an interaction term between the 2 predictors to assess the moderation effect, and adjusted for previous-day pain, menstrual status, education level, and body mass index. **Results**: The association of previous-day (day *t-1*) exercise to pain symptoms on day *t* was moderated by habitual exercise frequency, independent of covariates (Rate ratio=0.96, 95% CI=0.95, 0.98, p=0.0007 for day-level pain score, B=-0.14, 95%CI=-0.26, -0.016, p=0.026 for Δ-score). Those who regularly engaged in exercise at least 3 times per week were more likely to experience favorable pain outcomes after having a bout of exercise on the previous day. **Conclusions:** Regular exercise might influence the day-level (i.e., short-term) association of pain symptoms to exercise. These findings can inform exercise recommendations for endometriosis pain management, especially for those who are at greater risk for lack of regular exercise due to acute exacerbations in their pain after exercise.

Strengths and limitations of this study

- This study leverages data from a research mobile app (Phendo) designed for collecting self-reported data on symptoms and self-management of endometriosis.
- Daily exercise and pain symptom patterns in endometriosis is investigated under ecologically-valid conditions.
- The participant sample (N=1,009) represents 38 countries, ages across the reproductive life span, and various person-level characteristics.
- The study is limited by self-reported data collection by somewhat consistent trackers and lacks details on duration or intensity of exercise to evaluate as potential moderators.
- Participants consisted of mostly white, non-Hispanic white individuals; therefore, results might not be generalizable to other demographic groups.

INTRODUCTION

Exercise, a subset of physical activity (PA) that is planned, structured, repetitive, and intended to improve or maintain physical fitness, is an important component of effective pain management (i.e., reduction and prevention of pain symptoms).[1, 2] Both acute (i.e., single bout/session) and chronic (i.e., repeated bouts/sessions over time) exercise training have been demonstrated to improve numerous pain-related conditions. [1, 3-7] However, pain-related responses to exercise are variable in populations with chronic pain conditions.[8] Similarly, exacerbation of pain with exercise could pose a barrier to regular exercise in such individuals, thus increasing resistance to exercising, which in return can worsen pain, related disability, and risk for co-morbidities.[9-11] Investigation into the naturally-occurring pattern of pain symptoms associated with exercise behavior can help inform the design of exercise-based therapies for targeting disease-related pain symptoms.

Individuals with endometriosis may benefit from such investigations for several reasons.[12-14] Endometriosis is a systemic, estrogen-dependent inflammatory condition characterized primarily by chronic pelvic and abdominal pain, pain with sexual intercourse, and infertility.[15, 16] It significantly impacts daily function and quality of life (QoL)[17, 18], contributing to a productivity loss of 6.3 hours/week[19] and an estimated $69.4 billion in excess health expenditures annually in the United States.[20] Existing medical and hormonal therapies have limited efficacy for pain management, often confounded by side effects.[21] Opioids and other analgesics are commonly prescribed for long-term use,[22, 23] despite treatment guidelines recommending use of nonpharmacologic therapies including PA.[24] Consequently, there is a critical need to identify alternative approaches for endometriosis pain management.

One such approach is exercise, based on various mechanisms proposed in the literature[25] that might pertain to endometriosis. These include regulation of the serotonergic and opioid receptors,[26] reduction of inflammatory markers associated with pain,[27, 28] and effect of exercise on nerve growth factor expression that is associated with the painful endometriosis lesions.[29, 30] Exercise can increase pain management self-efficacy, which is associated with improved pain outcomes and QoL, for individuals with chronic pain.[31] While the evidence on exercise for pain management is promising [4, 32, 33], existing data are scarce, cross-sectional, and indicate variable effects on pain outcomes.[33-37] Despite these limitations, previous reports of exercise-induced adaptations to pain stimuli through increased pain threshold suggest that the regularity with which an individual engages in exercise over the long term (i.e., habitual exercise frequency) might influence (i.e., moderate) the relationship between their day-level exercise and pain symptoms.[38, 39] Among regular exercisers, pain-related activation has been demonstrated in the brain's descending antinociceptive pathway, with corresponding reductions in self-reported pain after acute bouts of at least moderate intensity exercise.[40] Moreover, studies report that habitual exercise frequency moderates a variety of self-reported outcomes (e.g., mood, anxiety, fatigue) in response to acute exercise.[41-43] While these findings are promising, their generalizability is limited by sample characteristics, laboratory-based experimental pain stimuli and exercise manipulations, and brief measurement duration of up to several hours. Thus, further investigation is needed to examine the relationship between pain symptoms and exercise behavior with a representative sample, under ecologically valid conditions, while accounting for possible between-individual variability and temporal lags in the outcome that extend beyond several hours.

Accordingly, this study examines the naturally-occurring daily patterns of pain symptoms and exercise behavior in endometriosis. We leverage mobile self-tracking, a particularly useful approach for capturing ecologically valid profiles of the dynamic temporal fluctuations and between-individual variability in pain over time.[44] We primarily aim to delineate the degree to which an individual's typical weekly exercise frequency (i.e., habitual exercise) influences (i.e., moderates) the association of their pain symptoms on a given day (day *t*) to their previous-day (day *t-1*) exercise behavior (i.e., lagged-day effects). Given the previously documented variable course of pain symptomology in endometriosis,[45] we also delineate the variability in day-to-day pain experiences within these analyses.

MATERIALS AND METHODS

Study Design

Study design and protocols were approved by the Columbia University Irving Medical Center (CUIMC) Institutional Review Board (#AAAQ9812). This study was conducted with retrospective data collected through the observational research mobile app "Phendo". Phendo was designed and developed for self-tracking endometriosis symptoms and its management. It is available for iOS[1] and Android[2] in App stores for free.

Study Sample and Inclusion Criteria

The study sample comprised Phendo users with a self-reported surgery-, clinician-, or suspected diagnosis of endometriosis and self-tracked exercise and pain data between November 2016 and April 2020. All participants, regardless of diagnosis type, are provided the same set of measures for completion in the App. In a previous study, the endometriosis phenotype (i.e., characterization) obtained using Phendo data was consistent with both the characterization of the

disease in the literature based on standard clinical surveys and clinician (i.e., human expert) evaluations.[46] We decided a priori to include all participants who selected one of the three affirmative responses in the present analyses, excluding those who indicated that they did not have endometriosis. Out of the initial eligible pool of 9,792 Phendo users with reported endometriosis, 7,949 had at least one day of tracking of the variables of interest for the study. Of these, 1,009 users had sufficient amount of data on pain and exercise for analysis (See *Data Analysis*) and were included in the study.

Recruitment and Informed Consent

Study participants were passively recruited through one of the App stores, engagement on study social media sites, or word-of-mouth. Upon downloading Phendo, all potential users went through an informed consent and enrollment process before tracking any data. First, they were provided with an explanation of the App, its overall purpose and link to its website (citizenendo.org) which includes detailed information and instructional videos for using the App. Participants completed a brief "verify your understanding" quiz to ensure their comprehension of how their data might be used for research purposes, anonymity and confidentiality (See Supplementary Figures 1-2 for example screenshots). This was followed by formal electronic informed consent (and assent for individuals 13-18 years old), a copy of which was sent to the participant. Once enrolled, users were instructed to track daily, but they were free to track as much or as sporadically as they wished, and they did not receive any prompts or requests to track a specific variable from the research team. Findings from a previous study evaluating recruitment and retention patterns within Phendo and seven other similar self-tracking apps indicated that Phendo's user engagement was similar to standard engagement patterns in research smartphone

apps.[47] Participants in the current study did not receive financial compensation for their tracking activities.

Study Measures

Day-level Pain. We assessed day-level pain through multiple items within Phendo: 1. "Are you in pain now? Where is the pain?", 2. "Any gastrointestinal or urinary issues?" (painful urination (dysuria), painful bowel movement (dyschezia)). Phendo pain item response options include all areas of the body (20 available choices, as well as right/left and upper/middle/lower specification), and can be mapped onto a visual, analogous to the McGill Pain Scale.[48] Pain severity for each affirmative response was rated on a 3-point categorical scale (mild, moderate, or severe), analogous to other commonly used pain rating scales in the literature.[49, 50] This categorization has been used for standardization and comparisons across different pain measures, and demonstrated superior ability to capture the nonlinear relationship between reported pain severity and interference with activity than use of numbers.[51, 52]

We computed a heuristic, composite day-level pain score to capture participants' conceptualization of their pain experience by summing the severity scores reported for each body area (e.g., moderate pain in abdomen, mild pains in chest and leg would yield *2+1+1=4* as the total score).[45] This allowed consideration of the multi-dimensional pain experience in a single outcome. To account for and circumvent any potential pain rumination/catastrophizing [53, 54] and varying tracking habits among participants, the score was computed based on the unique reports of area-severity pairs per day for each participant (e.g., if a participant tracked mild abdominal pain three times in a day, this abdomen-mild pair is counted toward the daily pain score only once). This score was the foundation of two study outcome variables: 1) total day-level pain score, and 2) difference in day-level pain score from previous day to the next (i.e., *t*-(*t*-

1)). The latter captures additional nuances in the data, enabling analyses to distinguish between participants with overall high day-level pain scores over time and experience a post-exercise reduction in pain versus those with low pain scores and who not experience a post-exercise reduction in pain. In the current study sample, the composite pain scores were moderately correlated with scores from other standard pain measures (e.g., $r=0.36$, $p<0.0001$ with the Pelvic-Abdominal Pain Visual Analog Scale (VAS); $r=-0.46$, $p<0.0001$ with Medical Outcomes Study 36-item Health Survey (SF-36) Bodily Pain subscale).

Day-level and habitual exercise. Phendo allows tracking of daily exercise through responding to a root question "Did you exercise today? (Yes/No)". Upon selecting a "Yes", users can further customize their entry within this item by adding exercise details through unrestricted free-text responses. We used responses to the root item to compute day-level and mean weekly exercise frequency (i.e., habitual exercise) for each participant. We calculated the latter by summing the number of exercise reports tracked per week across the range of days of data and then dividing this number by the total number of weeks of data. We used free-text responses to categorize exercises by modality and to validate that the entries were exercise-related. Any non-exercise activity (e.g., sleep, meditate, sitting, socialize) was recoded as a no exercise in the analytic data set. This day-level exercise assessment aims to increase ecological validity[55, 56] and reduce the likelihood of low test-retest reliability and inaccuracy due to recall bias.[57] We evaluated the validity of the scores from the Phendo exercise item through a series of analyses with the study sample.[58] Results supported its concurrency with other self-reported recall-based measures (i.e., Kendall's $\tau=0.256$, $p<0.001$ with Exercise Vital Sign[59] and $\tau=0.294$, $p=0.001$ with accelerometers; $B=18.73$, $p=0.039$ in association to the Nurses' Health Study II Weekly Exercise Scale[60] scores).

Standard Pain and Exercise Measures. To allow comparisons of the study sample with others in the literature, we report sample summary scores from the following components of the World Endometriosis Research Foundation (WERF) Endometriosis Patient Questionnaire (EPQ-S)[61, 62]: 1) The 2-item Bodily Pain subscale of the SF-36,[63] 2) Pelvic-abdominal Pain VAS ("*Please rate how severe your general pelvic/lower abdominal pain was at its worst in the last 3 months using the pain scale below where 0=no pain and 10=worst imaginable pain*."), and 3) The 8-item Nurses' Health Study II Weekly Physical Activity Scale (NHS-II) [60]. It measures self-reported weekly durations of major exercise modalities (i.e., walking, running, lap swimming, jogging, bicycling, tennis, calisthenics, other aerobic recreation) in a typical week in the past 12 months. These durations can further be multiplied by their metabolic equivalents (METs) based on the Compendium of PA [64] and summed to obtain the total weekly exercise-related energy expenditure (EE). We report both the total weekly minutes and EE for the sample.

Patient and Public Involvement

We developed Phendo measures using patient-centered participatory design, through qualitative (focus groups, interviews) and quantitative research (surveys, coded content analysis) with participants with endometriosis, described in detail elsewhere.[65, 66] This technique for developing patient-reported outcome measures has been suggested to enhance content validity and relevance of the measure to the target population, thus providing a more comprehensive and accurate representation of the disease under study.[55, 67-69]

Data Analysis

Sample Characteristics. We characterized the study sample through frequencies (%) and means (standard deviation; SD) of demographics, self-reported pain medication use, and scores on the standard pain and exercise measures for those who completed the surveys. We

characterized pain symptomology in the sample by describing the prevalence of self-tracked pain severities by each body area.

Associations of pain symptoms with exercise behavior. Using generalized linear mixed models (GLMMs), we separately estimated day-level total pain score and pain score difference as primary outcomes. Both outcomes were regressed on previous-day (day *t-1*) exercise and mean weekly exercise frequency to estimate the slope of mean pain level on day *t* and change in pain. We included an interaction term between the 2 predictors to assess the moderation of the day-level association by each individual's mean weekly exercise frequency. We included participant as a random effect to account for between-person variability in daily pain by estimating a separate intercept for each participant. Models were further adjusted for menstrual status (binary: yes/no), previous-day (i.e., day *t-1*) pain, body mass index (BMI) and education level. Race/ethnicity and age were not significantly associated with average daily pain reports ($F=1.68$, $p=0.14$ for race/ethnicity; $r=-0.148$, $p=0.07$ for age), and age was further significantly associated with education level (Kruskal-Wallis $X^2=64.948$, $p<0.0001$). To avoid redundancy and multicollinearity, race/ethnicity and age were not included as model covariates.

Model Specification. We specified a zero-inflated negative binomial (ZINB) distribution when modeling the total pain outcome, as it has been demonstrated to provide the best fit for outcomes with over-dispersion and zero-inflation (i.e., zeros due to both sampling and missingness) [70-72]. Missing values in the BMI (22%), education level (19%) and menstrual status (22%) were imputed as described in Supplementary File 1 and checked for appropriateness based on convergence and marginal distributions following guidelines [73-75] (See Supplementary Figures 3-5). Adequacy of imputations for valid statistical inference were verified based on the recommended measures of missing data information of *fraction of missing*

information (λ) and *relative increase in variance due to nonresponse* (r)[76, 77] (See Supplementary Table 1). Further details of the model specification are in Supplementary file 1. We included participants who had at least 11 pairs of consecutive days of data in the final analytic sample as this provided sufficient amount of data to 1) ensure model convergence and improve reliability and accuracy of the estimates, particularly the random effects and their variances[78-81], and 2) adequately infer participants' habitual exercise level by considering at least three weeks' worth of tracking to compute the weekly exercise frequency. Finally, as a post-hoc analysis, we tested the possible influence of type of endometriosis diagnosis by including this categorical variable in the 2 models described above. We conducted the data analyses using R[82] and the glmmTMB package for the GLMMs.[71, 72] Statistical significance level was set at $p<0.05$ for all analyses.

RESULTS

Sample Characteristics. Sample characteristics are provided in Table 1. Participants (N=1,009) had on average 89.6 days of data available for analysis (SD=62.8, Range=22-841, IQR=31). Participants collectively represented 38 countries, with a wide age range (14-63 years), and varying education and employment status. Almost 70% (N=702) had laparoscopic confirmation of their diagnosis, 19.8% (N=200) had a clinician diagnosis, and 10.6% (N=107) had suspected endometriosis (i.e., "I think I have endometriosis (know the symptoms, no doctor)"). Scores from the VAS, SF-36, and NHS-II Scales are provided in Table 2. The overall prevalence of non-prescription pain medication use, opioid-based medication use, opioid-paracetamol/acetaminophen combination medication use were 49.35%, 11.19%, and 11.39%, respectively (See Table 1).

Pain symptom patterns. Mean daily pain score was 4.48 (SD=7.11, 0-79). Mean person-level daily pain score (i.e., "mean of means") was 4.82 (SD=4.57, Range=0-34). Moderate intensity was the most frequently reported severity across all body areas (Mean=49.3%, SD=22.2), and pelvic pain was the most prevalent area, followed by back pain and gastrointestinal pain (See Figure 1).

Habitual exercise patterns. Mean weekly exercise frequency was 1.43/week (SD=1.54, Range=0-6.87/week, IQR=2.21). The exercise frequencies were at least 3 times per week 21.3% (N=215); 1-2 times per week, 40.2% (N=406); and no regular exercise, 38.5% (388). Prevalence of the 10 most frequently reported exercise modalities in the sample are depicted in Figure 2. Walking was the most common modality, reported by 50.94% of the participants, followed by yoga (30.82%), and muscle strength/endurance training activities (24.38%). Yoga and stretching exercises were collectively reported by almost 45% of the sample.

Association of day-level pain to exercise. Tables 3 and 4 display results of the GLMMs estimating day-level total pain score and difference. Coefficients for the model interaction terms indicated a small but statistically significant moderation of previous-day exercise by habitual exercise frequency (RR=0.96 for total pain score and -0.14 for pain score difference, $p<0.05$; See Figure 3). Further inspection of this interaction indicated a mean typical exercise frequency of ~3 times/week as the point after which previous-day exercise began to be associated with favorable pain outcomes (e.g., a decrease from the predicted mean score) on the following day, adjusted for other day-level and person-level factors (Figure 3). On the other hand, those who exercised less frequently or none were more likely to report higher levels of pain and larger increases (or smaller decreases) in pain 1 day after an exercise bout compared to not having exercised the day before.

Variability in estimated pain scores. There was substantial between-person variability in average day-level pain scores, based on the statistically significant random effect of participant in the models (See Tables 3 and 4, also depicted in Figure 4). We quantified the significance of this random effect through a restricted likelihood ratio test (RLRT) based on simulations from the model sample distribution.[83, 84] This yielded an observed likelihood ratio (RLRT =7183.3, p-value < 0.0001), indicating substantial contribution of the random effect to the total model pain variance.

Post-hoc analyses. Inclusion of diagnosis type in the model did not have an influence on the results based on the non-significant B coefficients (p=0.48 and p=0.59 for pain score and p=0.70 and p=0.27 for difference in pain score). There were no differences across the 3 groups with respect to either daily total pain score or difference (χ^2 = 1415.1, df = 1438, p-value = 0.661) (See Supplementary Tables 2 and 3 for full results).

DISCUSSION

Summary of findings. We leveraged 90,382 days of mHealth self-tracking data from 1,009 women with endometriosis to investigate the association between exercise behavior and day-level fluctuations in pain. For the average individual, the association between previous-day exercise to pain was moderated by their habitual exercise frequency, i.e., the frequency with which they engaged in exercised in a typical week. This effect was consistent across participants and independent of person-level covariates. There was substantial between-person heterogeneity in day-level pain patterns. To our knowledge, this is the first study to quantify the association between day-level pain symptoms and exercise in an international sample of women with

endometriosis and to identify habitual weekly exercise frequency as a moderator of this relationship.

Moderation effects. Previous-day exercise was associated with more favorable pain outcomes for participants who engaged in regular exercise at least 3 times per week in our sample. That is, these participants were more likely to report lower pain score and smaller increases (or larger decreases) in pain the day after an exercise bout, compared to not having exercised the previous day. In contrast, those who engaged in regular exercise less than twice a week were more likely to experience pain symptoms on days after having engaged in exercise. This is in line with the physical activity guidelines [85, 86], which recommend aerobic exercise at least 3 times per week and muscle-strengthening exercise at least twice per week.[87] However, there are no specific recommendations for endometriosis in the current guidelines; and systematic reviews recommend "overall, general exercise" without further details due to lack adequate research on the optimal dose of exercise for endometriosis pain.[4, 36] Our findings provide preliminary evidence for informing exercise recommendations for endometriosis pain management (i.e., prevention or reduction), specifically for targeting those who are at greater risk for insufficient regular exercise due to acute exacerbations in their pain after exercise. This moderation effect suggests that an individual might need to develop a regular, sustained exercise behavior (i.e., habit) to start experiencing the favorable pain outcomes associated with acute bouts of exercise. Nevertheless, future experimental studies are warranted for a comprehensive investigation of this question.

Patterns of pain symptoms. Our findings of moderate pain in pelvis as the most frequently reported pain are in line with those from others on endometriosis[88] and various chronic pain conditions.[89, 90] The distribution of the total daily pain scores was right-skewed

(i.e., extreme scores on the higher ends of the range) with a mean score that was on the lower end of the range. This could partly be due to the data collection method which includes not just days where the participant experienced pain but also days without pain. Indeed, our participants on average did not report or experience any pain 6.25% of the time. In contrast, traditional study designs typically rely on recall of past pain experience aggregated over a period of time (e.g., past week, month) and ask the participant to report their average or highest pain severity over this period.[91, 92] Such recall-based techniques are prone to peak-and-end effects,[93] and catastrophizing or other similar biases.[92, 94] Recruitment from clinical referral points is a common practice and this has been attributed to higher normative scores in the literature,[91] as opposed to more even distributions of pain symptomology among community-based samples.[95] Self-tracking facilitates documentation of not only severe pain, but also mild, moderate, and no pain instances, therefore enabling a more realistic representation of the pain experience as it dynamically unfolds over time. This can reduce the likelihood of over-representing severe cases, which is a potential limitation attributed to data collected at point of contact in clinical settings.[17] However, it is difficult to make direct comparisons with other studies given the different pain measures, warranting further research.

Patterns of exercise behavior. The mean weekly exercise frequency in the study sample was 1.43/week (SD=1.57, IQR=2.29), with only 24.5 % (N=202) engaging in exercise at least three times a week. This suggests that individuals with endometriosis might be at increased risk for physical inactivity[85, 87], which is a risk factor for various comorbidities [96] and further linked to exacerbation of chronic pain.[97, 98] These collectively underscore the need to focus efforts on promoting regular exercise in women with endometriosis. Notably, yoga and stretching were reported collectively by almost half of the sample within Phendo. This could

indicate that participants use these approaches for pain relief, in line with a previous study reporting efficacy of hatha yoga.[33] Nevertheless, participants overall tracked a wide range of exercise modalities across the intensity spectrum (e.g., yoga vs running/cycling) as helpful for their symptoms, suggesting between-individual variability in responses to a given exercise type or intensity. This can be targeted through individualized exercise prescriptions,[25, 99] providing precedence for undertaking a precision approach for pain self-management in endometriosis. Various individualization approaches (e.g., adaptive treatment strategies,[100] micro-randomized trials,[101] just-in-time adaptive interventions [102]) have been investigated for intervening on health behaviors, including PA.[5, 101] It would be opportune to implement a similar N-of-1 intervention approach for identifying person-specific optimal "dose" of exercise based on its parameters to target endometriosis pain symptoms.

Consideration of person-level factors. Another novel finding in our study was the similar point estimates for the effect of exercise on pain outcomes between those with clinician/surgical- versus suspected diagnosis of endometriosis. Endometriosis is difficult to diagnose, with a 7.6-year delay between symptom onset and its surgical diagnosis.[20, 103, 104] Endometriosis patients further face insurance-related challenges in accessing healthcare for their condition.[15, 105] The participants without a formal diagnosis might have sought medical care for their symptoms but not received the needed care (e.g., diagnostic testing, referral to a specialist), received false negative diagnostic tests results,[103] or lacked adequate access to healthcare. This finding underscores the need for further research in endometriosis that considers self-report of endometriosis symptoms, instead of limiting to patients with a physician referral or relying on secondary data sources (e.g., electronic health records).

Novel methodological contributions. In contrast to other existing questionnaires in the literature, the self-tracking items in Phendo measure momentary and daily pain symptoms and exercise –a time interval for which there are no standard validated, commonly used measures designed for frequent sampling. Phendo's pain tracking items are similar in design to other pain measures,[48, 66] and have been indicated to be reflective of pain documentation in clinical records.[45] While mHealth studies have examined the validity, utility and specificity for various pain conditions [52, 106, 107] of their pain measurement approaches, a standard "all-in-one" single outcome that captures the multi-dimensional pain experience across different populations remains to be established.[53, 108]Computation of a composite pain has been proposed by others[109] as this circumvents numerous limitations in current pain assessment approaches, including lack of a standard single outcome that can be used universally,[108] or a validated instrument that captures all the constructs of persistent pain.[110] There is furthermore a lack of endometriosis-specific pain measures for repeated assessments, thus the heuristic composite pain measure allowed consideration of two dimensions of pain simultaneously in our analyses. The pain scores in the current study sample were moderately correlated with those from the pelvic-abdominal VAS and the SF-36 bodily pain measure, which were also similarly correlated with each other ($r=0.46$, $p<0.0001$). Nevertheless, future directions include evaluation of this measure in larger samples for its reliability and validity via a nomological network-based analysis.

Limitations. We acknowledge several limitations of this study, including reliance on self-reports for the type of endometriosis diagnosis and exercise behavior. First, we used a binary measure of exercise in our analyses and did not have sufficient details on duration or intensity for inclusion in the analyses as potential moderators. Of note, similar mHealth measures of daily PA and exercise have been used by others [111-113] who reported concordance with accelerometer-

based measures,[114] and higher correlations than self-report methods with accelerometer measures.[111, 112] While we provide preliminary evidence toward the validity of Phendo's exercise tracking item both as a day-level and habitual measure[58], future studies are needed to evaluate it in larger samples and compare against research-grade accelerometers. Similarly, we did not have granular daily data on pain medication use, as such it was not investigated as a potential covariate in the analyses. In addition to medications, future studies could consider other pain management approaches for comparison to exercise, given previous research suggesting endometriosis patients report using a variety of symptom management techniques.[45] Next, our sample consisted primarily of White, non-Hispanic women who are relatively consistent mHealth technology users and furthermore can understand English to use the App. Therefore, the results might differ among other groups including non-English speakers or those without an interest in mHealth use for self-management or monitoring.

Conclusion

In this study, we provide evidence that habitual exercise frequency is a potential moderator of the association between pain symptoms and previous-day exercise in endometriosis, indicating that those who regularly exercise at least ~3 times per week are less likely to report pain symptoms after having exercised on the previous day. Individuals with endometriosis are significantly more likely to have higher all-cause healthcare utilization and direct health care costs than those without endometriosis, including twice the prevalence of opioid prescriptions for pain management [23] and prolonged duration of prescriptions.[22] While guidelines recommend prescribing exercise for management of pain in clinical populations, endometriosis (or general chronic) pain-specific recommendations to guide patients

and providers on measurable parameters (time, type, intensity, and frequency) are lacking. Future studies are warranted investigating the effects of both acute and chronic exercise on endometriosis pain with a focus on various types, intensities and durations.

Author Contributions

IE conceptualized the study, conducted the data analyses, and prepared the first draft of the manuscript. SLG and ENH were responsible for data acquisition, curation and management. NE acquired the funding and provided the mHealth infrastructure for the study (Phendo App). NE and SB provided guidance on the study design and data analyses. SB, NE, SLG and ENH critically reviewed and provided feedback on the manuscript.

Funding

Funding for the work is provided by a postdoctoral fellowship from the Data Science Institute at Columbia University and an award from the National Library of Medicine (R01 LM013043). We are grateful to the Phendo participants.

Competing Interests

All authors report no conflicts of interest.

Data availability statement

Data are available on reasonable request.

1. Available at https://itunes.apple.com/us/app/phendo/id1145512423
2. Available at https://play.google.com/store/apps/details?id=com.appliedinformaticsinc.phendo

References

1. Ambrose KR, Golightly YM. Physical exercise as non-pharmacological treatment of chronic pain: Why and when. Best Pract Res Clin Rheumatol. 2015;29(1):120-30. Epub 2015/05/23. doi: 10.1016/j.berh.2015.04.022. PubMed PMID: 26267006.

2. Rice D, Nijs J, Kosek E, Wideman T, Hasenbring MI, Koltyn K, et al. Exercise-Induced Hypoalgesia in Pain-Free and Chronic Pain Populations: State of the Art and Future Directions. The Journal of Pain. 2019;20(11):1249-66. doi: https://doi.org/10.1016/j.jpain.2019.03.005.

3. Lemmens J, De Pauw J, Van Soom T, Michiels S, Versijpt J, van Breda E, et al. The effect of aerobic exercise on the number of migraine days, duration and pain intensity in migraine: a systematic literature review and meta-analysis. J Headache Pain. 2019;20(1):16. Epub 2019/02/16. doi: 10.1186/s10194-019-0961-8. PubMed PMID: 30764753; PubMed Central PMCID: PMCPMC6734345.

4. Armour M, Ee CC, Naidoo D, Ayati Z, Chalmers KJ, Steel KA, et al. Exercise for dysmenorrhoea. Cochrane Database of Systematic Reviews. 2019;(9). doi: 10.1002/14651858.CD004142.pub4. PubMed PMID: CD004142.

5. Rabbi M, Aung MS, Gay G, Reid MC, Choudhury T. Feasibility and acceptability of mobile phone–based auto-personalized physical activity recommendations for chronic pain self-management: pilot study on adults. Journal of medical Internet research. 2018;20(10):e10147.

6. Sevel L, Boissoneault J, Alappattu M, Bishop M, Robinson M. Training endogenous pain modulation: a preliminary investigation of neural adaptation following repeated exposure to clinically-relevant pain. Brain Imaging and Behavior. 2020;14(3):881-96. doi: 10.1007/s11682-018-0033-8.

7. Gordon R, Bloxham S, editors. A systematic review of the effects of exercise and physical activity on non-specific chronic low back pain. Healthcare; 2016: Multidisciplinary Digital Publishing Institute.

8. Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH. Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. Cochrane Database of Systematic Reviews. 2017;(4).

9. Zhang R, Chomistek AK, Dimitrakoff JD, Giovannucci EL, Willett WC, Rosner BA, et al. Physical activity and chronic prostatitis/chronic pelvic pain syndrome. Med Sci Sports Exerc. 2015;47(4):757-64. doi: 10.1249/MSS.0000000000000472. PubMed PMID: 25116086.

10. Pinto A, Di Raimondo D, Tuttolomondo A, Buttà C, Milio G, Licata G. Effects of physical exercise on inflammatory markers of atherosclerosis. Current pharmaceutical design. 2012;18(28):4326-49.

11. Garatachea N, Molinero O, Martínez-García R, Jimenez-Jimenez R, Gonzalez-Gallego J, Marquez S. Feelings of well being in elderly people: relationship to physical activity and physical function. Archives of Gerontology and Geriatrics. 2009;48(3):306-12.

12. Tennfjord MK, Gabrielsen R, Tellum T. Effect of physical activity and exercise on endometriosis-associated symptoms: a systematic review. BMC Women's Health. 2021;21(1):355. doi: 10.1186/s12905-021-01500-4.

13. Evans S, Fernandez S, Olive L, Payne LA, Mikocka-Walus A. Psychological and mind-body interventions for endometriosis: A systematic review. J Psychosom Res. 2019;124:109756. Epub 2019/08/25. doi: 10.1016/j.jpsychores.2019.109756. PubMed PMID: 31443810.

14. Mira TAA, Buen MM, Borges MG, Yela DA, Benetti-Pinto CL. Systematic review and meta-analysis of complementary treatments for women with symptomatic endometriosis. Int J Gynaecol Obstet. 2018;143(1):2-9. Epub 2018/06/27. doi: 10.1002/ijgo.12576. PubMed PMID: 29944729.

15. Fourquet J, Gao X, Zavala D, Orengo JC, Abac S, Ruiz A, et al. Patients' report on how endometriosis affects health, work, and daily life. Fertility and sterility. 2010;93(7):2424-8. doi: 10.1016/j.fertnstert.2009.09.017. PubMed PMID: 19926084.

16. Schliep KC, Mumford SL, Peterson CM, Chen Z, Johnstone EB, Sharp HT, et al. Pain typology and incident endometriosis. Hum Reprod. 2015;30(10):2427-38. Epub 2015/08/11. doi: 10.1093/humrep/dev147. PubMed PMID: 26269529.

17. De Graaff A, D'hooghe T, Dunselman G, Dirksen C, Hummelshoj L, Consortium WE, et al. The significant effect of endometriosis on physical, mental and social wellbeing: results from an international cross-sectional survey. Human reproduction. 2013;28(10):2677-85.

18. Simoens S, Dunselman G, Dirksen C, Hummelshoj L, Bokor A, Brandes I, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. Human Reproduction. 2012;27(5):1292-9.

19. Soliman AM, Coyne KS, Gries KS, Castelli-Haley J, Snabes MC, Surrey ES. The effect of endometriosis symptoms on absenteeism and presenteeism in the workplace and at home. Journal of managed care & specialty pharmacy. 2017;23(7):745-54.

20. Simoens S, Dunselman G, Dirksen C, Hummelshoj L, Bokor A, Brandes I, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. Hum Reprod. 2012;27(5):1292-9. Epub 2012/03/17. doi: 10.1093/humrep/des073. PubMed PMID: 22422778.

21. The Practice Committee of the American Society for Reproductive Medicine. Treatment of pelvic pain associated with endometriosis: a committee opinion. 2014 2014/04/01/. Report No.: 0015-0282 Contract No.: 4.

22. Lamvu G, Soliman AM, Manthena SR, Gordon K, Knight J, Taylor HS. Patterns of prescription opioid use in women with endometriosis: evaluating prolonged use, daily dose, and concomitant use with benzodiazepines. Obstetrics and gynecology. 2019;133(6):1120.

23. Soliman AM, Surrey ES, Bonafede M, Nelson JK, Vora JB, Agarwal SK. Health care utilization and costs associated with endometriosis among women with medicaid insurance. Journal of managed care & specialty pharmacy. 2019;25(5):566-72.

24. Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain—United States, 2016. Jama. 2016;315(15):1624-45.

25. Sluka KA, Frey-Law L, Hoeger Bement M. Exercise-induced pain and analgesia? Underlying mechanisms and clinical translation. Pain. 2018;159 Suppl 1(Suppl 1):S91-S7. doi: 10.1097/j.pain.0000000000001235. PubMed PMID: 30113953.

26. Tour J, Löfgren M, Mannerkorpi K, Gerdle B, Larsson A, Palstam A, et al. Gene-to-gene interactions regulate endogenous pain modulation in fibromyalgia patients and healthy controls-antagonistic effects between opioid and serotonin-related genes. Pain. 2017;158(7):1194-203. Epub 2017/03/11. doi: 10.1097/j.pain.0000000000000896. PubMed PMID: 28282362; PubMed Central PMCID: PMCPMC5472004.

27. Bobinski F, Teixeira JM, Sluka KA, Santos ARS. Interleukin-4 mediates the analgesia produced by low-intensity exercise in mice with neuropathic pain. Pain. 2018;159(3):437-50. Epub 2017/11/16. doi: 10.1097/j.pain.0000000000001109. PubMed PMID: 29140923; PubMed Central PMCID: PMCPMC5812806.

28. Montenegro ML, Bonocher CM, Meola J, Portella RL, Ribeiro-Silva A, Brunaldi MO, et al. Effect of Physical Exercise on Endometriosis Experimentally Induced in Rats. Reproductive sciences (Thousand Oaks, Calif). 2018:1933719118799205. Epub 2018/09/21. doi: 10.1177/1933719118799205. PubMed PMID: 30231769.

29. Stratton P, Berkley KJ. Chronic pelvic pain and endometriosis: translational evidence of the relationship and implications. Human reproduction update. 2011;17(3):327-46. Epub 2010/11/23. doi: 10.1093/humupd/dmq050. PubMed PMID: 21106492.

30. Park S-J, Yong M-S, Na S-S. Effect of exercise on the expression of nerve growth factor in the spinal cord of rats with induced osteoarthritis. Journal of physical therapy science. 2015;27(8):2551-4. Epub 2015/08/21. doi: 10.1589/jpts.27.2551. PubMed PMID: 26357438.

31. Karasawa Y, Yamada K, Iseki M, Yamaguchi M, Murakami Y, Tamagawa T, et al. Association between change in self-efficacy and reduction in disability among patients with chronic pain. PLOS ONE. 2019;14(4):e0215404. doi: 10.1371/journal.pone.0215404.

32. Armour M, Sinclair J, Chalmers KJ, Smith CA. Self-management strategies amongst Australian women with endometriosis: a national online survey. BMC Complementary and Alternative Medicine. 2019;19(1):17. doi: 10.1186/s12906-019-2431-x.

33. Gonçalves AV, Barros NF, Bahamondes L. The Practice of Hatha Yoga for the Treatment of Pain Associated with Endometriosis. Journal of Alternative & Complementary Medicine. 2017;23(1):45-52. doi: 10.1089/acm.2015.0343. PubMed PMID: 120746246.

34. Ricci E, Viganò P, Cipriani S, Chiaffarino F, Bianchi S, Rebonato G, et al. Physical activity and endometriosis risk in women with infertility or pain: Systematic review and meta-analysis. Medicine. 2016;95(40):e4957-e. doi: 10.1097/MD.0000000000004957. PubMed PMID: 27749551.

35. Carpenter SE, Tjaden B, Rock JA, Kimball A. The effect of regular exercise on women receiving danazol for treatment of endometriosis. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics. 1995;49(3):299-304. Epub 1995/06/01. PubMed PMID: 9764869.

36. Bonocher CM, Montenegro ML, Rosa ESJC, Ferriani RA, Meola J. Endometriosis and physical exercises: a systematic review. Reprod Biol Endocrinol. 2014;12:4. Epub 2014/01/08. doi: 10.1186/1477-7827-12-4. PubMed PMID: 24393293; PubMed Central PMCID: PMCPMC3895811.

37. Naugle KM, Fillingim RB, Riley JL, 3rd. A meta-analytic review of the hypoalgesic effects of exercise. The journal of pain : official journal of the American Pain Society. 2012;13(12):1139-50. Epub 2012/11/08. doi: 10.1016/j.jpain.2012.09.006. PubMed PMID: 23141188.

38. Janal MN, Colt EWD, Clark WC, Glusman M. Pain sensitivity, mood and plasma endocrine levels in man following long-distance running: Effects of naloxone. Pain. 1984;19(1):13-25. doi: https://doi.org/10.1016/0304-3959(84)90061-7.

39. Droste C, Greenlee MW, Schreck M, Roskamm H. Experimental pain thresholds and plasma beta-endorphin levels during exercise. Medicine & Science in Sports & Exercise. 1991;23(3):334-42. doi: 10.1249/00005768-199103000-00012.

40. Scheef L, Jankowski J, Daamen M, Weyer G, Klingenberg M, Renner J, et al. An fMRI study on the acute effects of exercise on pain processing in trained athletes. PAIN. 2012;153(8).

41. Hoffman MD, Hoffman DR. Exercisers Achieve Greater Acute Exercise-Induced Mood Enhancement Than Nonexercisers. Archives of Physical Medicine and Rehabilitation. 2008;89(2):358-63. doi: 10.1016/j.apmr.2007.09.026.

42. Hallgren M, Moss ND, Gastin P. Regular exercise participation mediates the affective response to acute bouts of vigorous exercise. J Sports Sci Med. 2010;9(4):629-37. Epub 2010/01/01. PubMed PMID: 24149790; PubMed Central PMCID: PMCPMC3761821.

43. Chen Y-C, Chen C, Martínez RM, Etnier JL, Cheng Y. Habitual physical activity mediates the acute exercise-induced modulation of anxiety-related amygdala functional connectivity. Scientific Reports. 2019;9(1):19787. doi: 10.1038/s41598-019-56226-z.

44. May M, Junghaenel DU, Ono M, Stone AA, Schneider S. Ecological Momentary Assessment Methodology in Chronic Pain Research: A Systematic Review. J Pain. 2018;19(7):699-716. Epub 2018/01/31. doi: 10.1016/j.jpain.2018.01.006. PubMed PMID: 29371113.

45. Ensari I, Pichon A, Lipsky-Gorman S, Bakken S, Elhadad N. Augmenting the Clinical Data Sources for Enigmatic Diseases: A Cross-Sectional Study of Self-Tracking Data and Clinical Documentation in Endometriosis. Applied Clinical Informatics. 2020;11(05):769-84.

46. Urteaga I, McKillop M, Elhadad N. Learning endometriosis phenotypes from patient-generated data. npj Digital Medicine. 2020;3(1):88. doi: 10.1038/s41746-020-0292-9.

47. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. npj Digital Medicine. 2020;3(1):21. doi: 10.1038/s41746-020-0224-8.

48. Melzack R. The McGill Pain Questionnaire: Major properties and scoring methods. PAIN. 1975;1(3).

49. Jones KR, Vojir CP, Hutt E, Fink R. Determining mild, moderate, and severe pain equivalency across pain-intensity tools in nursing home residents. J Rehabil Res Dev. 2007;44(2):305-14. Epub 2007/06/07. doi: 10.1682/jrrd.2006.05.0051. PubMed PMID: 17551881.

50. Bestel E, Gotteland J-P, Donnez J, Taylor RN, Garner EI. Linzagolix for Endometriosis-Associated Pain: Lipid Changes After 52 Weeks of Treatment [25B]. Obstetrics & Gynecology. 2020;135:25S. doi: 10.1097/01.AOG.0000663180.46470.c9. PubMed PMID: 00006250-202005001-00082.

51. Serlin RC, Mendoza TR, Nakamura Y, Edwards KR, Cleeland CS. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. Pain. 1995;61(2):277-84. Epub 1995/05/01. doi: 10.1016/0304-3959(94)00178-h. PubMed PMID: 7659438.

52. Adams P, Murnane EL, Elfenbein M, Wethington E, Gay G. Supporting the Self-Management of Chronic Pain Conditions with Tailored Momentary Self-Assessments. Proc SIGCHI Conf Hum Factor Comput Syst. 2017;2017:1065-77. doi: 10.1145/3025453.3025832. PubMed PMID: 30310887.

53. Boonstra AM, Stewart RE, Köke AJA, Oosterwijk RFA, Swaan JL, Schreurs KMG, et al. Cut-Off Points for Mild, Moderate, and Severe Pain on the Numeric Rating Scale for Pain in Patients with Chronic Musculoskeletal Pain: Variability and Influence of Sex and Catastrophizing. Front Psychol. 2016;7:1466-. doi: 10.3389/fpsyg.2016.01466. PubMed PMID: 27746750.

54. Dirks JF, Wunder J, Kinsman R, McElhinny J, Jones NF. A Pain Rating Scale and a Pain Behavior Checklist for Clinical Use: Development, Norms, and the Consistency Score. Psychotherapy and Psychosomatics. 1993;59(1):41-9.

55. Frost MH, Reeve BB, Liepa AM, Stauffer JW, Hays RD. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? Value Health. 2007;10 Suppl 2:S94-s105. Epub 2007/11/13. doi: 10.1111/j.1524-4733.2007.00272.x. PubMed PMID: 17995479.

56. Faurholt-Jepsen M, Munkholm K, Frost M, Bardram JE, Kessing LV. Electronic self-monitoring of mood using IT platforms in adult patients with bipolar disorder: A systematic review of the validity and evidence. BMC Psychiatry. 2016;16(1):7. doi: 10.1186/s12888-016-0713-0.

57. Charter RA. Sample size requirements for precise estimates of reliability, generalizability, and validity coefficients. J Clin Exp Neuropsychol. 1999;21(4):559-66. Epub 1999/11/07. doi: 10.1076/jcen.21.4.559.889. PubMed PMID: 10550813.

58. Ensari I, Horan E, Elhadad N, Bakken S. Evaluation of a disease-specific mHealth-based exercise self-tracking measure. MedRXiv. 2022. doi: https://doi.org/10.1101/2022.05.16.22275170.

59. Kuntz JL, Young DR, Saelens BE, Frank LD, Meenan RT, Dickerson JF, et al. Validity of the Exercise Vital Sign Tool to Assess Physical Activity. Am J Prev Med. 2021;60(6):866-72. Epub 2021/03/31. doi: 10.1016/j.amepre.2021.01.012. PubMed PMID: 33781618; PubMed Central PMCID: PMCPMC8154650.

60. Wolf AM, Hunter DJ, Colditz GA, Manson JE, Stampfer MJ, Corsano KA, et al. Reproducibility and validity of a self-administered physical activity questionnaire. Int J Epidemiol. 1994;23(5):991-9. Epub 1994/10/01. doi: 10.1093/ije/23.5.991. PubMed PMID: 7860180.

61. Vitonis AF, Vincent K, Rahmioglu N, Fassbender A, Buck Louis GM, Hummelshoj L, et al. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonization Project: II. Clinical and covariate phenotype data collection in endometriosis research. Fertility and sterility. 2014;102(5):1223-32. Epub 2014/09/22. doi: 10.1016/j.fertnstert.2014.07.1244. PubMed PMID: 25256930.

62. Jones G, Kennedy S, Barnard A, Wong J, Jenkinson C. Development of an endometriosis quality-of-life instrument: The Endometriosis Health Profile-30. Obstet Gynecol. 2001;98(2):258-64. Epub 2001/08/17. doi: 10.1016/s0029-7844(01)01433-8. PubMed PMID: 11506842.

63. McHorney CA, Ware JE, Lu JFR, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36): III. Tests of Data Quality, Scaling Assumptions, and Reliability across Diverse Patient Groups. Medical Care. 1994;32(1):40-66.

64. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc. 2011;43(8):1575-81.

65. McKillop M, Voigt N, Schnall R, Elhadad N. Exploring self-tracking as a participatory research activity among women with endometriosis. Journal of Participatory Medicine. 2016.

66. McKillop M, Mamykina L, Elhadad N, editors. Designing in the Dark: Eliciting Self-tracking Dimensions for Understanding Enigmatic Disease. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; 2018: ACM.

67. Anthoine E, Moret L, Regnault A, Sébille V, Hardouin J-B. Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. Health Qual Life Outcomes. 2014;12:176-. doi: 10.1186/s12955-014-0176-2. PubMed PMID: 25492701.

68. US Department of Health Human Services. Guidance for industry-Patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.

69. Lomas J, Pickard L, Mohide A. Patient versus clinician item generation for quality-of-life measures: the case of language-disabled adults. Medical Care. 1987:764-9.

70. Hu M-C, Pavlicova M, Nunes EV. Zero-inflated and hurdle models of count data with extra zeros: examples from an HIV-risk reduction intervention trial. Am J Drug Alcohol Abuse. 2011;37(5):367-75. doi: 10.3109/00952990.2011.597280. PubMed PMID: 21854279.

71. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R journal. 2017;9(2):378-400.

72. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. Modeling zero-inflated count data with glmmTMB. bioRxiv. 2017:132753. doi: 10.1101/132753.

73. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software; Vol 1, Issue 3 (2011). 2011.

74. Van Buuren S. Flexible imputation of missing data: CRC press; 2018.

75. Bondarenko I, Raghunathan T. Graphical and numerical diagnostic tools to assess suitability of multiple imputations and imputation models. Statistics in Medicine. 2016;35(17):3007-20. doi: https://doi.org/10.1002/sim.6926.

76. Rubin DB. The Calculation of Posterior Distributions by Data Augmentation: Comment: A Noniterative Sampling/Importance Resampling Alternative to the Data Augmentation Algorithm for Creating a Few Imputations When Fractions of Missing Information Are Modest: The SIR Algorithm. Journal of the American Statistical Association. 1987;82(398):543-6. doi: 10.2307/2289460.

77. Rubin DB. Multiple imputation for nonresponse in surveys: John Wiley & Sons; 2004.

78. Schunck R. Cluster Size and Aggregated Level 2 Variables in Multilevel Models. A Cautionary Note. 2016. 2016;10(1). Epub 2016-07-20. doi: 10.12758/mda.2016.005.

79. Bell B, Ferron J, Kromrey J, editors. Cluster Size in Multilevel Models: The Impact of Sparse Data Structures on Point and Interval Estimates in Two-Level Models2008.

80. Austin PC, Leckie G. The effect of number of clusters and cluster size on statistical power and Type I error rates when testing random effects variance components in multilevel linear and logistic regression models. Journal of Statistical Computation and Simulation. 2018;88(16):3151-63. doi: 10.1080/00949655.2018.1504945.

81. Snijders TAB. Power and sample size in multilevel modeling. In: Everitt B, Howell D, editors. Encyclopedia of Statistics in Behavioral Science. 3: Wiley; 2006. p. 1570–3.

82. Core Team R. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria1997.

83. Scheipl F, Greven S, Küchenhoff H. Size and power of tests for a zero random effect variance or polynomial regression in additive and linear mixed models. Computational statistics & data analysis. 2008;52(7):3283-99.

84. Crainiceanu CM, Ruppert D. Likelihood ratio tests in linear mixed models with one variance component. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2004;66(1):165-85.

85. US Department of Health Human Services Physical activity guidelines advisory committee scientific report. 2018.

86. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med. 2020;54(24):1451-62. doi: 10.1136/bjsports-2020-102955. PubMed PMID: 33239350; PubMed Central PMCID: PMCPMC7719906.

87. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. Jama. 2018;320(19):2020-8.

88. Warzecha D, Szymusik I, Wielgos M, Pietrzak B. The Impact of Endometriosis on the Quality of Life and the Incidence of Depression-A Cohort Study. Int J Environ Res Public Health. 2020;17(10). Epub 2020/05/28. doi: 10.3390/ijerph17103641. PubMed PMID: 32455821; PubMed Central PMCID: PMCPMC7277332.

89. Becker N, Thomsen AB, Olsen AK, Sjogren P, Bech P, Eriksen J. Pain epidemiology and health related quality of life in chronic non-malignant pain patients referred to a Danish multidisciplinary pain center. Pain. 1997;73(3):393-400. doi: 10.1016/s0304-3959(97)00126-7. PubMed PMID: WOS:000071429200014.

90. Bouhassira D, Lantéri-Minet M, Attal N, Laurent B, Touboul C. Prevalence of chronic pain with neuropathic characteristics in the general population. Pain. 2008;136(3):380-7.

91. Nicholas MK, Asghari A, Blyth FM. What do the numbers mean? Normative data in chronic pain measures. Pain. 2008;134(1):158-73. doi: https://doi.org/10.1016/j.pain.2007.04.007.

92. Dansie EJ, Turk DC. Assessment of patients with chronic pain. Br J Anaesth. 2013;111(1):19-25. doi: 10.1093/bja/aet124. PubMed PMID: 23794641.

93. Schneider S, Stone AA, Schwartz JE, Broderick JE. Peak and end effects in patients' daily recall of pain and fatigue: a within-subjects analysis. J Pain. 2011;12(2):228-35. doi: 10.1016/j.jpain.2010.07.001. PubMed PMID: 20817615.

94. De Boer M, Struys M, Versteegen G. Pain-related catastrophizing in pain patients and people with pain in the general population. European journal of pain. 2012;16(7):1044-52.

95. Ehde DM, Gibbons LE, Chwastiak L, Bombardier CH, Sullivan MD, Kraft GH. Chronic pain in a large community sample of persons with multiple sclerosis. Multiple Sclerosis Journal. 2003;9(6):605-11. doi: 10.1191/1352458503ms939oa.

96. Katzmarzyk PT, Powell KE, Jakicic JM, Troiano RP, Piercy K, Tennant B, et al. Sedentary Behavior and Health: Update from the 2018 Physical Activity Guidelines Advisory Committee. Med Sci Sports Exerc. 2019;51(6):1227-41. doi: 10.1249/MSS.0000000000001935. PubMed PMID: 31095080.

97. Basen-Engquist K, Scruggs S, Jhingran A, Bodurka DC, Lu K, Ramondetta L, et al. Physical activity and obesity in endometrial cancer survivors: associations with pain, fatigue, and physical functioning. Am J Obstet Gynecol. 2009;200(3):288.e1-.e2888. Epub 2008/12/25. doi: 10.1016/j.ajog.2008.10.010. PubMed PMID: 19110220.

98. Dansie EJ, Turk DC, Martin KR, Van Domelen DR, Patel KV. Association of Chronic Widespread Pain With Objectively Measured Physical Activity in Adults: Findings From the National Health and Nutrition Examination Survey. The Journal of Pain. 2014;15(5):507-15. doi: https://doi.org/10.1016/j.jpain.2014.01.489.

99. Polaski AM, Phelps AL, Kostek MC, Szucs KA, Kolber BJ. Exercise-induced hypoalgesia: A meta-analysis of exercise dosing for the treatment of chronic pain. PloS one. 2019;14(1):e0210418-e. doi: 10.1371/journal.pone.0210418. PubMed PMID: 30625201.

100. Almirall D, Compton SN, Gunlicks-Stoessel M, Duan N, Murphy SA. Designing a pilot sequential multiple assignment randomized trial for developing an adaptive treatment strategy. Statistics in medicine. 2012;31(17):1887-902. Epub 2012/03/23. doi: 10.1002/sim.4512. PubMed PMID: 22438190; PubMed Central PMCID: PMCPMC3399974.

101. Klasnja P, Smith S, Seewald NJ, Lee A, Hall K, Luers B, et al. Efficacy of Contextually Tailored Suggestions for Physical Activity: A Micro-randomized Optimization Trial of HeartSteps. Annals of behavioral medicine : a publication of the Society of Behavioral Medicine. 2018. Epub 2018/09/08. doi: 10.1093/abm/kay067. PubMed PMID: 30192907.

102. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. Annals of behavioral medicine : a publication of the Society of Behavioral Medicine. 2018;52(6):446-62. Epub 2016/09/25. doi: 10.1007/s12160-016-9830-8. PubMed PMID: 27663578; PubMed Central PMCID: PMCPMC5364076.

103. Falcone T, Mascha E. The elusive diagnostic test for endometriosis. Fertility and sterility. 2003;80(4):886-8.

104. Marian S, Hermanowicz-Szamatowicz K. Endometriosis–a decade later–still an enigmatic disease. What is the new in the diagnosis and treatment? Gynecological Endocrinology. 2020;36(2):104-8.

105. Fourquet J, Zavala DE, Missmer S, Bracero N, Romaguera J, Flores I. Disparities in healthcare services in women with endometriosis with public vs private health insurance. American Journal of Obstetrics & Gynecology. 2019;221(6):623.e1-.e11. doi: 10.1016/j.ajog.2019.06.020.

106. Lee RR, Rashid A, Ghio D, Thomson W, Cordingley L. "Seeing Pain Differently": A Qualitative Investigation Into the Differences and Similarities of Pain and Rheumatology Specialists' Interpretation of Multidimensional Mobile Health Pain Data From Children and Young People With Juvenile Idiopathic Arthritis. JMIR Mhealth Uhealth. 2019;7(7):e12952. Epub 2019/07/04. doi: 10.2196/12952. PubMed PMID: 31267979; PubMed Central PMCID: PMCPMC6632104.

107. Jamison RN, Raymond SA, Levine JG, Slawsby EA, Nedeljkovic SS, Katz NP. Electronic diaries for monitoring chronic pain: 1-year validation study. Pain. 2001;91(3):277-85.

108. Bouhassira D, Attal N. All in one: Is it possible to assess all dimensions of any pain with a simple questionnaire? PAIN. 2009;144(1).

109. Pilitsis JG, Fahey M, Custozzo A, Chakravarthy K, Capobianco R. Composite score is a better reflection of patient response to chronic pain therapy compared with pain intensity alone. Neuromodulation: Technology at the Neural Interface. 2021;24(1):68-75.

110. Grimmer-Somers K, Vipond N, Kumar S, Hall G. A review and critique of assessment instruments for patients with persistent pain. Journal of pain research. 2009;2:21.

111. Knell G, Gabriel KP, Businelle MS, Shuval K, Wetter DW, Kendzor DE. Ecological Momentary Assessment of Physical Activity: Validation Study. J Med Internet Res. 2017;19(7):e253. doi: 10.2196/jmir.7602.

112. Swendeman D, Comulada WS, Koussa M, Worthman CM, Estrin D, Rotheram-Borus MJ, et al. Longitudinal Validity and Reliability of Brief Smartphone Self-Monitoring of Diet, Stress, and Physical Activity in a Diverse Sample of Mothers. JMIR Mhealth Uhealth. 2018;6(9):e176. Epub 2018/09/27. doi: 10.2196/mhealth.9378. PubMed PMID: 30249576; PubMed Central PMCID: PMCPMC6231816.

113. Katapally TR, Chu LM. Digital epidemiological and citizen science methodology to capture prospective physical activity in free-living conditions: a SMART Platform study. BMJ Open. 2020;10(6):e036787. Epub 2020/07/01. doi: 10.1136/bmjopen-2020-036787. PubMed PMID: 32595163; PubMed Central PMCID: PMCPMC7322321.

114. Zink J, Belcher BR, Dzubur E, Ke W, O'Connor S, Huh J, et al. Association Between Self-Reported and Objective Activity Levels by Demographic Factors: Ecological Momentary Assessment Study in Children. JMIR Mhealth Uhealth. 2018;6(6):e150. doi: 10.2196/mhealth.9592.

Table 1. Study Sample Characteristics.

Characteristic (N)	Mean (SD) / Frequency (%)
Age (827)	31.0 (7.26), Median=30.6 (MAD=7.41), Range= 14.3-62.9
BMI (787)	25.9 (6.98), Median=24.1 (MAD=4.74), Range= 16.01-72.24
Type of endometriosis diagnosis	
Surgery (702)	69.57 %
Clinician (200)	19.82 %
Self-diagnosis (107)	10.60 %
Work Environment	
Home (218)	26.42 %
Outside (570)	69.09 %
Unknown (221)	21.29 %
Living environment	
Rural (129)	15.27 %
Suburban (340)	41.21 %
Urban (363)	44.00 %
Unknown (161)	19.5 %
Relationship status	
Married/domestic partnership (442)	53.57 %
Separated/divorced (28)	3.39 %
Single/never married (310)	37.57 %

Unknown (229)	22.69 %
Education Level	
College or higher (547)	66.30 %
High school graduate or less (74)	8.96 %
Some college (209)	25.33 %
Unknown (179)	17.7 %
Employment Status	
Employed (541)	65.57 %
Not employed (120)	14.54 %
Student (129)	15.63 %
Unknown (219)	21.70 %
Race/Ethnicity	
White, Non-Hispanic (699)	84.72 %
Black, Non-Hispanic (20)	2.42 %
Asian (22)	2.6 %
Native American (6)	0.72 %
Hispanic (38)	4.6 %
Other (51)	6.18 %
Unknown (173)	17.14 %
Country of Residence	
United States (444)	44.0 %
United Kingdom (83)	8.22 %
Canada (75)	7.43 %

Australia (59)	5.84 %
Germany (38)	3.76 %
New Zealand (34)	3.36 %
Other (69)	6.83 %
Unknown (207)	20.51 %

Table 2. Sample Study Scores on Standard Measures of Pain and Exercise.

EPQ-S Measures (N)	**Mean (SD)**
SF-36 Bodily Pain (375)	35.47 (22.33)
Pelvic-abdominal pain VAS (316)	7.37 (1.97)
NHS-II PA Scale Total Weekly Minutes (359)	175.2 (280.2)
NHS-II PA Scale Total Weekly EE (359)	16.13 (30.37)

Table 3. Results of the regression model estimating day-level total pain score (N=1,009).

Conditional Random Effects		**Variance (95% CI)**	
Participant		1.09 (0.98, 1.21)	
Conditional Fixed Effects	**Rate Ratio (95% CI)**	**Log Odds (SE)**	***z*-score**
Intercept	4.26*** (3.26, 5.56)	1.45*** (0.13)	10.82
Menstrual Status	1.29*** (1.25, 1.32)	0.25*** (0.01)	20.31
Previous Day Pain	1.02*** (1.02, 1.03)	0.02*** (0.00)	29.69
Body Mass Index	1.01* (1.00, 1.02)	0.01 (0.00)	2.02
Mean Weekly Exercise Frequency	0.93* (0.89, 0.97)	-0.06** (0.02)	-2.96
Previous Day exercise	1.10* (1.05, 1.15)	0.09**(0.15)	3.88
Some College Education Level	0.87 (0.83, 1.56)	0.13 (0.15)	0.86
College or Higher Education Level	0.93 (0.66, 1.16)	-0.13 (0.14)	-0.92
Mean Weekly Exercise Frequency * Previous Day exercise	0.96** (0.95, 0.98)	-0.03** (0.01)	-3.37
Zero Inflation Terms	**Rate Ratio (95% CI)**	**Log Odds (SE)**	***z*-score**
Intercept	0.17 (0.16, 0.18)	-1.73***(0.02)	-62.96
Same Day Exercise	5.34 (5.01, 5.68)	1.67*** (0.03)	52.53

95% CI=95% Confidence Interval. *p<0.05 ** p <0.001, ***p<0.0001. Previous day pain and BMI were sample mean-centered. BMI and education level were kept as covariates in the model based on their significant associations with mean day-level pain scores (Pearson's r=0.15 for BMI and Kruskal-Wallis $\chi^2 = 18.061$ for education level, p<0.001).

Table 4. Results of the regression model estimating pain score difference (N=1,009).

Conditional Random Effects	**Variance (95% CI)**
Participant (Intercept)	9.16 (8.28, 10.13)
Residual	26.83

Conditional Fixed Effects	**B coefficient (SE)**	**95% CI**	***z*-score**
Intercept	2.70*** (0.51)	1.68, 3.72	5.29
Menstrual Status	1.47*** (0.09)	1.28, 1.66	15.43
Previous Day Pain	-0.86*** (0.01)	-0.87, -0.85	-143.43
Body Mass Index	0.05* (0.01)	0.01, 0.10	2.86
Mean Weekly Exercise Frequency	-0.27** (0.08)	-0.44, -0.10	-3.12
Previous Day Exercise	0.92** (0.18)	0.56, 1.27	5.08
Some College Education Level	-0.84 (0.62)	-2.11, 0.42	-1.35
College or Higher Education Level	-2.07** (0.52)	-3.10, -1.03	-3.96
Mean Weekly Exercise Frequency * Previous Day Exercise	-0.14* (0.06)	-0.26, -0.01	-2.22

Zero Inflation Terms	**B coefficient**	**95% CI**	***z*-score**
Intercept	-0.91*** (0.01)	-0.93, -0.88	-63.84
Same Day Exercise	0.70*** (0.02)	0.66, 0.75	32.09

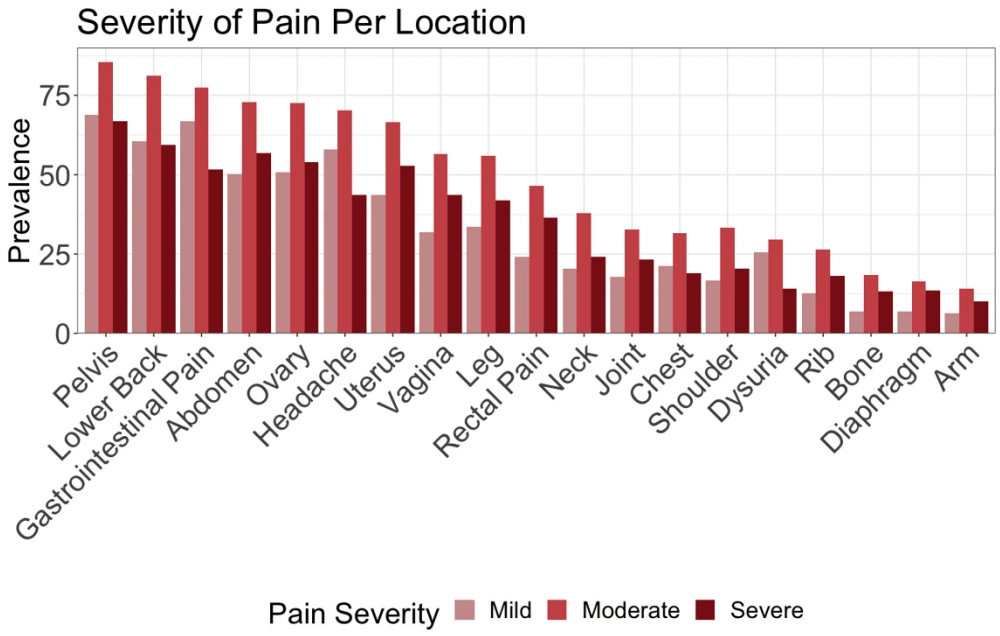
SE= Standard Error. *p<0.05 ** p <0.001, ***p<0.0001. Previous day pain and BMI were sample mean-centered. BMI and education level were kept as covariates in the model based on their significant associations with mean day-level pain scores (Pearson's r=0.15 for BMI and Kruskal-Wallis χ^2 = 18.061 for education level, p<0.001).

Figure 1. Prevalence of pain severity by location reported among participants (i.e., unique counts of body area-severity per participant). Moderate intensity was the most frequently tracked across all body areas (14.1%-85.4%).

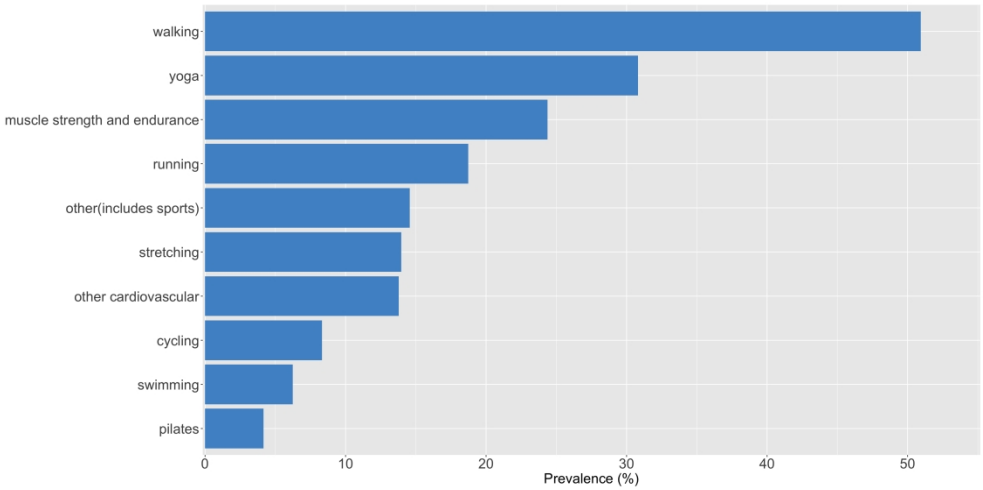
Figure 2. Prevalence of self-reported exercise modalities in the study sample. "Other cardiovascular" category includes activities such as dancing, aerobics and using the elliptical machine. "Muscle strength and endurance" category includes activities such as weight lifting and calisthenics. "Other exercise" category includes sports activities such as skiing and soccer, multi-modal exercises (e.g., high intensity interval training of both cardiovascular and muscular endurance), or those that did not fit into the other categories (e.g., stabilizing or balancing exercises, Wii fit or other home-based fitness activities).

Panel Figure 3. Moderation of effect of previous-day exercise by habitual exercise levels (X axes). Y axes represent predicted day-level total scores (top) and differences (bottom) in pain. Shaded areas depict 95% confidence intervals. At approximately 3 times/week of regular exercise, previous day exercise starts to be associated with more favorable pain outcomes on the following day (i.e., decrease from the model predicted mean scores), adjusted for other day-level and person level factors.

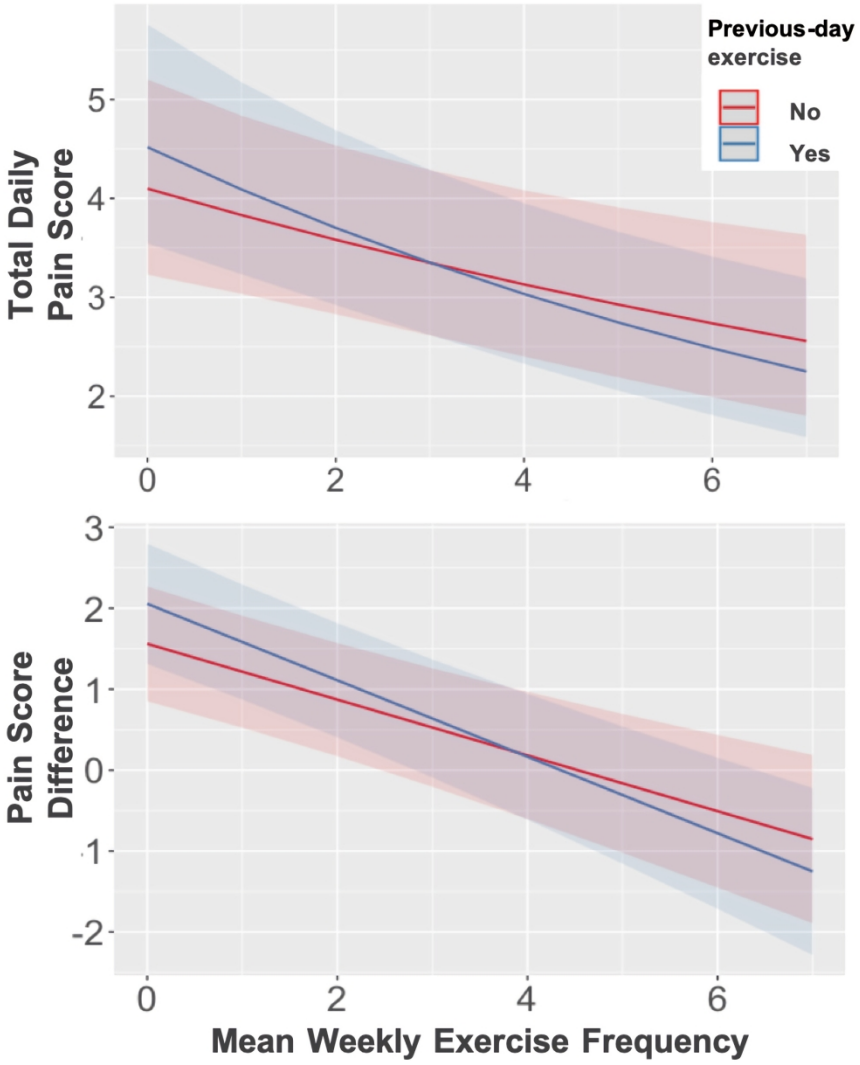
Figure 4. Plot of the random effect of the participant on total day pain scores estimated from the multilevel model (N=1,009). Y-axis represents the range of estimated average pain scores for each participant. Each black dot represents one participant's mean (i.e., random intercept), grey lines indicate 95% confidence intervals. Distribution of points across the x-axis indicate large variability across individuals (i.e., between-group variance), and the grey lines indicate the within-person variability in daily scores over time.



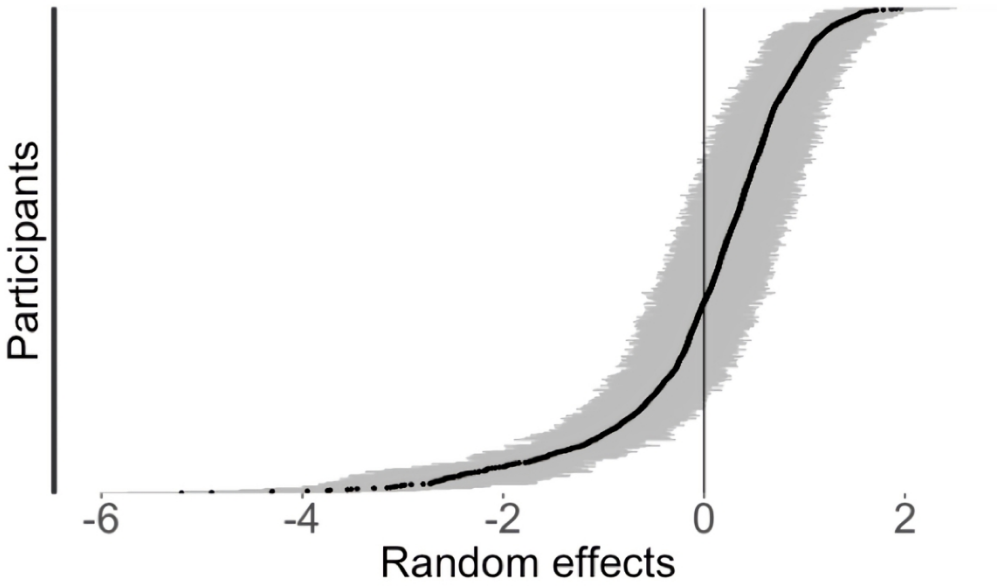
776x493mm (72 x 72 DPI)



1411x705mm (72 x 72 DPI)



694x750mm (72 x 72 DPI)



571x339mm (72 x 72 DPI)

Supplementary File 1. Missing Data Imputations.

Phendo is an observational research app and participants do not receive prompts from the research team to track any given item at a certain time. They are free to track (or not track) any given item as they wish. Consequently, missingness in the data occurs due to a variety of possible reasons that are not always known or easy to distinguish. For example, a period not tracked for a day could mean that the participant did not have a period, or they chose not to track, or did not use the app at all that day. To circumvent this issue, we took several measures. First, we limited data to days for which the participant tracked their pain, exercise and menstrual status at least once, as a proxy for app use. Next, we assigned a score of zero for pain on days where the participant had tracked exercise or menstrual status but not pain. This approach is motivated by 2 reasons. First, the nature of the pain question in Phendo (i.e., "Where is the pain?", "How severe is the pain?") assumes the participants to track when they feel pain and therefore a "No Pain" response is neither available in the app nor would make sense. Second, multiple imputation methods impute such that the resulting imputations are limited to the observed values and distributions. Thus by default it would omit the possibility of a zero in the resultant pain score distribution, which increases risk of overestimation of the scores in the sample.

BMI (calculated from participant reported height and weight) and education level were missing for 22% and 19% of the participants, respectively, and menstrual status was missing (i.e., not tracked) 22% of the time in the dataset. We imputed these 3 variables using multivariate imputations by chained equations [1] according to the heteroscedastic linear two-level structure of the data (i.e., hierarchical where, participant is the clustering variable) following standard multilevel multiple imputation methods. [1-4] We used two-level predictive mean matching for BMI and education level, which is a semi-parametric imputation method that limits imputations

to the observed values and can preserve non-linear relations in the observed data, therefore the imputations do not deviate from the observed distribution[5] and two-level logistic regression for imputing menstrual status, using the rest of the dataset as the predictors. As per published recommendations,[1, 2] we also included the raw pain variable (i.e., with the missing values) as a predictor, to account for the possibility of an association between the missingness pattern of pain to these imputed variables. To assess the plausibility of the imputations and any significant deviance from the structure of the raw, non-imputed data, we inspected the imputation convergence plots, distributions of the imputed variables which are provided in Supplementary Figures 3 and 4.

Model specification. We used a zero-inflated negative binomial (ZINB) distribution when modeling the total pain outcome, as it has been demonstrated to provide the best fit for outcomes with over-dispersion and zero-inflation.[6-8] ZINB models consider two sources of zero observations: "sampling zeros" that are part of the underlying sampling distribution (i.e., negative binomial) and "structural zeros" that cannot score anything other than zero (i.e., participant did not track).[6] This virtue of the ZINB models allows for specification of the imputed zeros and prevents the risk of over-estimating effects and generates more conservative estimates for predictors of interest by estimating a separate zero-inflation term, as well as conditional model.[6] We specified the zero-inflation term such that it was dependent on the exercise variable for the day, in addition to specifying an overall general zero-inflation structure in the outcome through inclusion of an intercept, based on recommendations. [8] Menstrual status was not a significant predictor of zero-inflation and therefore removed from the zero-inflation term during the modeling process. We included participants who had at least 11 pairs of consecutive days of data in the final analytic sample as this provided sufficient amount of data to

1) ensure model convergence and improve reliability and accuracy of the estimates, particularly the random effects and their variances[9-12], and 2) adequately infer participants' habitual weekly exercise frequency by considering at least three weeks' worth of tracking to compute the weekly exercise frequency.

Supplementary File 2. Imputation Model diagnostics.

Appropriateness and plausibility of the estimates from imputed models were inspected following published guidelines. First, we used measures of missing data information to assess pooled estimate variances. The fraction of missing information (λ) is interpreted as the proportion of variation in the parameter of interest due to the missing data. The relative increase in variance due to nonresponse (r) is interpreted as the proportional increase in the sampling variance of the parameter of interest that is due to the missing data. Values of λ over 0.5 indicate that the influence of the imputation model on the results is larger than that of the complete-data model, suggesting potential problems in the imputations. Supplementary Table 1 provides results of these variance estimates, indicating satisfactory imputation and model fit.

Supplementary Table 1. Measures of Missing data information

	Total Pain Score		**Difference in Pain**	
Conditional Fixed Effects	**λ**	***r***	**λ**	***r***
Intercept	0.21	0.27	0.23	0.31
Menstrual Status	0.13	0.15	0.19	0.23
Previous Day Pain	0.01	0.01	0.00	0.00
Body Mass Index	0.13	0.15	0.23	0.31
Mean Weekly Exercise Frequency	0.00	0.00	0.01	0.01
Previous Day exercise	0.01	0.01	0.00	0.00
Some College Education Level	0.26	0.36	0.35	0.55
College or Higher Education Level	0.23	0.31	0.21	0.28
Mean Weekly Exercise Frequency * Previous Day exercise	0.00	0.00	0.00	0.00
Zero Inflation Terms				
Intercept	0.00	0.00	0.00	0.00
Same Day Exercise	0.00	0.00	0.00	0.00

Next, we inspected propensity scores, which is a more recent and increasingly accepted method for inspecting the suitability of data imputation.[2, 13, 14] The goal is to compare the distributions of observed and imputed data conditional on the missingness probability. Under the missing at random (MAR) assumption, the conditional distributions of the observed and missing data should be similar if the assumed model for creating multiple imputations has a good fit. To do this, we first estimate the probability of each record being incomplete (i.e., "response propensity") in the presence of missing data by conditioning on the response indicators as well as the observed covariates. The probabilities are then averaged over the imputed datasets to obtain stability. Supplementary Figure 3 plots BMI, education category

and menstrual status against the propensity score in each dataset. The distributions of the blue and red points are match up well without significant discrepancies (e.g., mismatch in patterns, imputed data systematically shifted toward one side of the axis).

Supplementary Table 2. Post-hoc analyses with endometriosis diagnosis included as a covariate. Conditional model results of the negative binomial model estimation of day-level total pain score (N=608).

Random Effects	**Variance (95% CI)**	
Participant (Intercept)	1.10 (0.99, 1.22)	
Fixed Effects	**Log Odds (SE)**	**z-score**
Intercept	1.37*** (0.12)	10.97
Menstrual Status	0.25*** (0.01)	21.40
Previous day Pain	0.02*** (0.01)	21.40
Body Mass Index	0.01* (0.004)	2.81
Mean weekly Exercise Frequency	-0.06** (0.02)	-3.01
Previous day exercise	0.09** (0.02)	3.85
Clinician diagnosis of endometriosis	-0.07 (0.10)	0.01
Self-diagnosis of endometriosis	-0.11 (0.11)	-1.01
Some college education level	0.22 (0.13)	-1.63
College or higher education level	-0.01 (0.12)	-0.12
Mean weekly Exercise Frequency*Previous day exercise	-0.03*** (0.01)	-3.42

SE=Standard Error. *p=0.001, ** p <0.001, ***p<0.0001. B coefficients are rate ratios. BMI =Body Mass Index. BMI and previous day pain were group mean centered.

Supplementary Table 3. Post-hoc analyses with endometriosis diagnosis included as a covariate. Conditional model results of the regression model estimation of pain score difference (N=1009).

Conditional Random Effects	**Variance (95% CI)**
Participant (Intercept)	13.34 (12.09, 14.93)

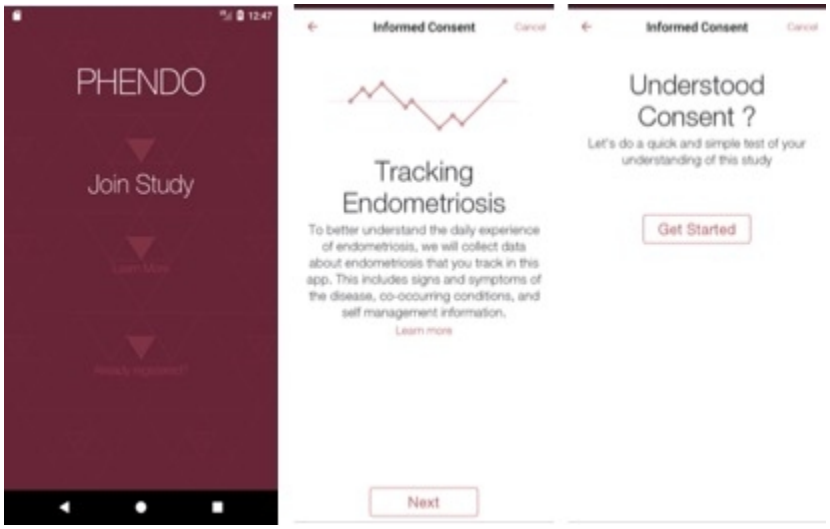
Fixed Effects	**B coefficient (SE)**	**z-score**
Intercept	2.45*** (0.46)	5.22
Menstrual status	1.46*** (0.08)	16.98
Previous day pain	-0.86*** (0.01)	-144.11
Body mass index	0.07* (0.01)	4.47
Mean weekly exercise frequency	-0.27** (0.09)	-3.03
Previous day exercise	0.92*** (0.18)	5.13
Clinician diagnosis of endometriosis	-0.05 (0.32)	-0.16
Self-diagnosis of endometriosis	-0.45 (0.43)	-1.29
Some college education level	-0.30 (0.51)	-0.58
College or higher education level	-1.72** (0.47)	-3.67
Mean weekly exercise frequency*Previous day exercise	-0.14* (0.06)	-2.31

SE=Standard Error. *p<0.05, ** p <0.01, ***p<0.0001. Body Mass Index and previous day pain were group mean centered.

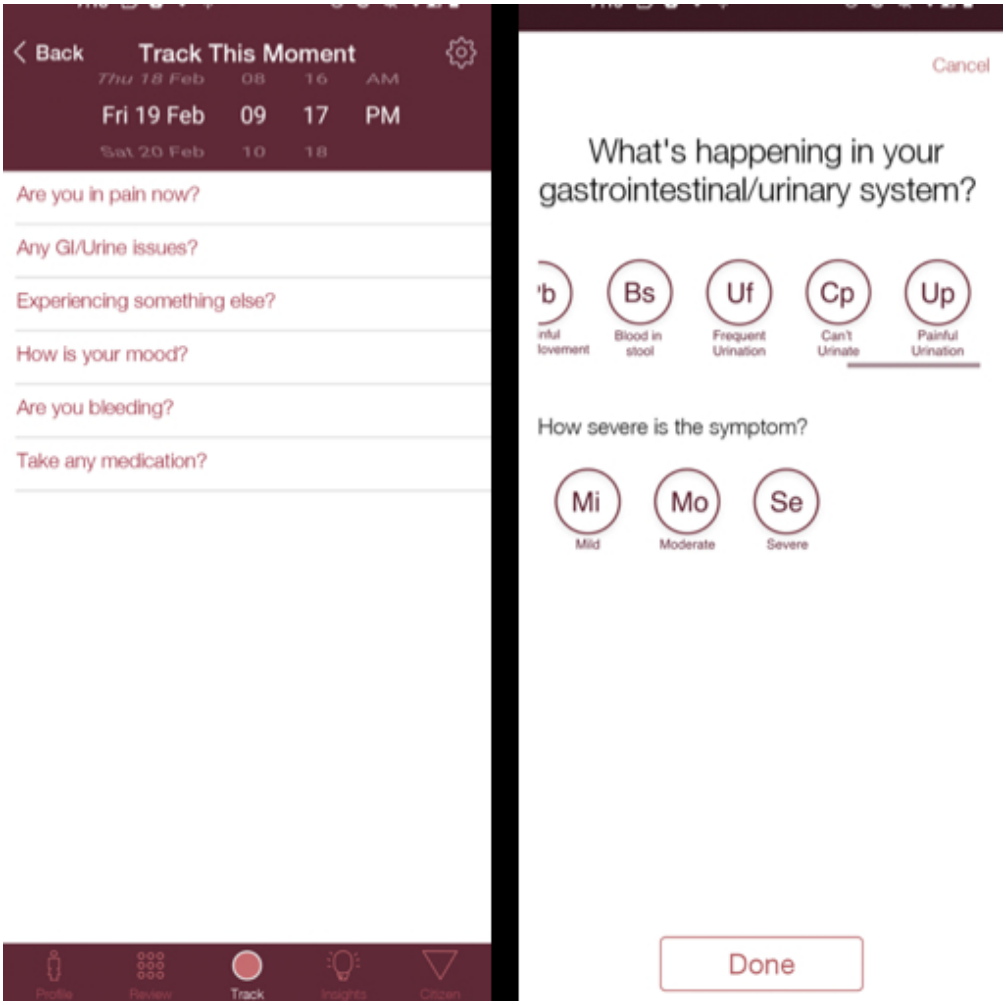
References

1. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software; Vol 1, Issue 3 (2011). 2011.
2. Van Buuren S. Flexible imputation of missing data: CRC press; 2018.
3. Grund S, Lüdtke O, Robitzsch A. Multiple imputation of missing data for multilevel models: Simulations and recommendations. Organizational Research Methods. 2018;21(1):111-49.
4. Barnard J, Rubin DB. Small-Sample Degrees of Freedom with Multiple Imputation. Biometrika. 1999;86(4):948-55.
5. Little RJ. Missing-data adjustments in large surveys. Journal of Business & Economic Statistics. 1988;6(3):287-96.

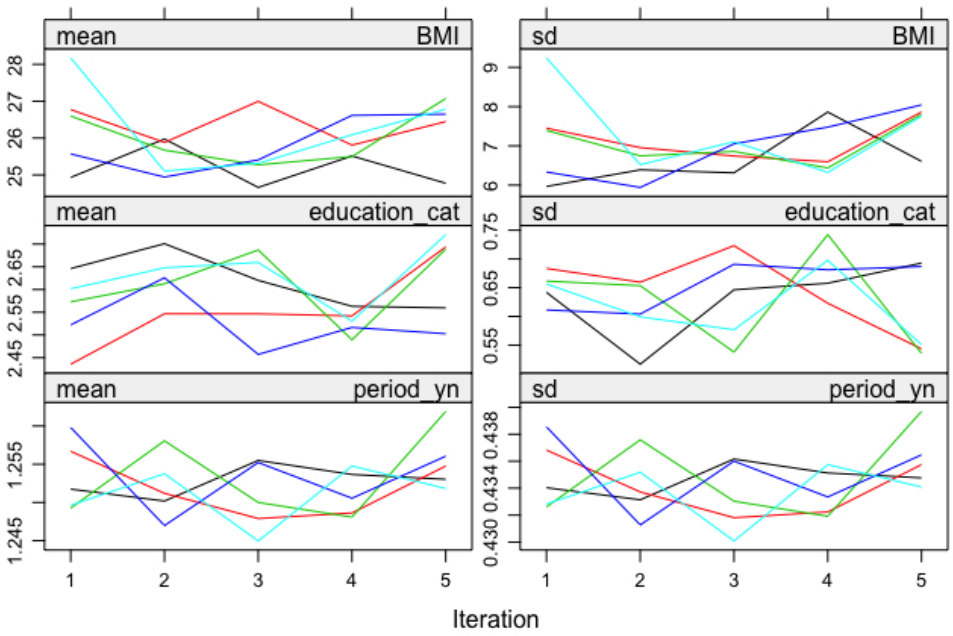
6. Hu M-C, Pavlicova M, Nunes EV. Zero-inflated and hurdle models of count data with extra zeros: examples from an HIV-risk reduction intervention trial. Am J Drug Alcohol Abuse. 2011;37(5):367-75. doi: 10.3109/00952990.2011.597280. PubMed PMID: 21854279.
7. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R journal. 2017;9(2):378-400.
8. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. Modeling zero-inflated count data with glmmTMB. bioRxiv. 2017:132753. doi: 10.1101/132753.
9. Schunck R. Cluster Size and Aggregated Level 2 Variables in Multilevel Models. A Cautionary Note. 2016. 2016;10(1). Epub 2016-07-20. doi: 10.12758/mda.2016.005.
10. Bell B, Ferron J, Kromrey J, editors. Cluster Size in Multilevel Models: The Impact of Sparse Data Structures on Point and Interval Estimates in Two-Level Models2008.
11. Austin PC, Leckie G. The effect of number of clusters and cluster size on statistical power and Type I error rates when testing random effects variance components in multilevel linear and logistic regression models. Journal of Statistical Computation and Simulation. 2018;88(16):3151-63. doi: 10.1080/00949655.2018.1504945.
12. Snijders TAB. Power and sample size in multilevel modeling. In: Everitt B, Howell D, editors. Encyclopedia of Statistics in Behavioral Science. 3: Wiley; 2006. p. 1570–3.
13. Nguyen CD, Carlin JB, Lee KJ. Model checking in multiple imputation: an overview and case study. Emerging Themes in Epidemiology. 2017;14(1):8. doi: 10.1186/s12982-017-0062-6.
14. Bondarenko I, Raghunathan T. Graphical and numerical diagnostic tools to assess suitability of multiple imputations and imputation models. Statistics in Medicine. 2016;35(17):3007-20. doi: https://doi.org/10.1002/sim.6926.



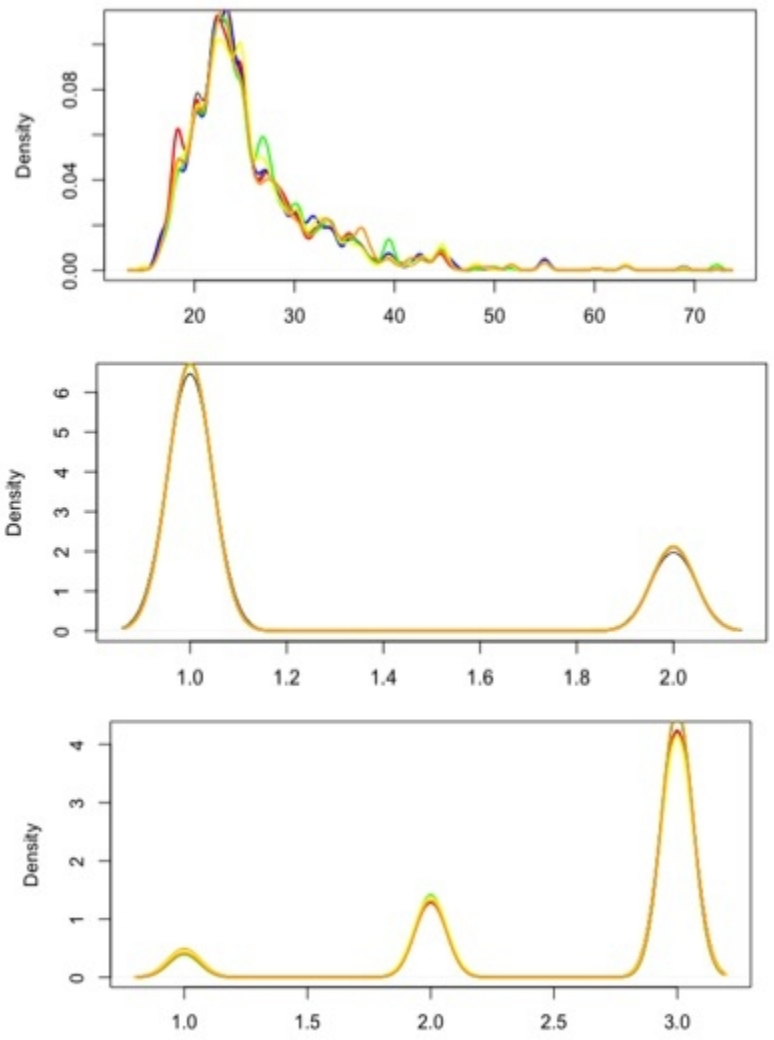
34x22mm (300 x 300 DPI)



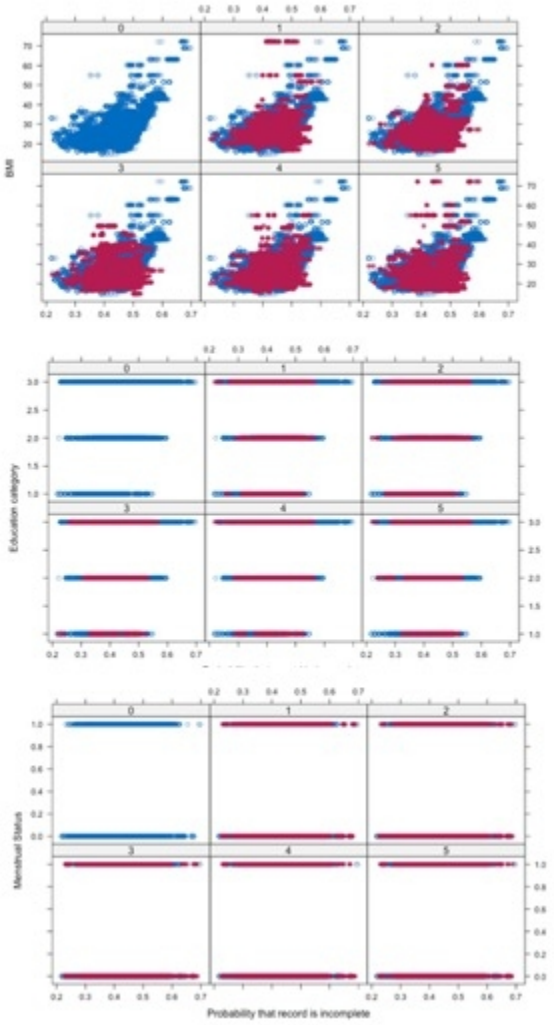
109x109mm (118 x 118 DPI)



142x99mm (118 x 118 DPI)



34x45mm (300 x 300 DPI)



23x44mm (300 x 300 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page # where this item is located:
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1,2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	6-7
Objectives	3	State specific objectives, including any prespecified hypotheses	8
Methods			
Study design	4	Present key elements of study design early in the paper	8-9
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	8-9
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	8-9
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed	N/A

		Case-control study—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	10-12
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	10-12
Bias	9	Describe any efforts to address potential sources of bias	10,11,14
Study size	10	Explain how the study size was arrived at	8,9

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	10-11
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	13-14
		(*b*) Describe any methods used to examine subgroups and interactions	13
		(*c*) Explain how missing data were addressed	13, Supplemental Files 1-2
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	13-14
		(*e*) Describe any sensitivity analyses	14, and Supplemental Tables
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	14
		(b) Give reasons for non-participation at each stage	9, 14
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	14, and Table 1
		(b) Indicate number of participants with missing data for each variable of interest	13-14, Table 1
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	N/A
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	*N/A*
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	*N/A*
		Cross-sectional study—Report numbers of outcome events or summary measures	15-16
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	15-16, Tables 2 and 3

	(*b*) Report category boundaries when continuous variables were categorized	14-15
	(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	15-16, and Supplemental Tables
Discussion			
Key results	18	Summarise key results with reference to study objectives	16-17
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	20-21
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	17-19
Generalisability	21	Discuss the generalisability (external validity) of the study results	20-21
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	22