Remote Perioperative Symptom Monitoring via Smartphone: A largescale feasibility analysis

Madelyn R. Frumkin[1], Gabrielle R. Messner[2], Katherine J. Holzer[3], Ziqi Xu[4,5], Thomas L. Rodebaugh[2], Haley Bernstein[3], Karen Frey[3], Saivee Ahuja[3], Chenyang Lu[4,5], & Simon Haroutounian[3,5]

[1]Center for Technology and Behavioral Health, Departments of Biomedical Data Science and Psychiatry, Geisel School of Medicine at Dartmouth, Lebanon, New Hampshire, USA
[2]Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[3]Department of Anesthesiology, Washington University School of Medicine in St. Louis, St. Louis, Missouri, USA
[4]Department of Computer Science & Engineering, Washington University School McKelvey School of Engineering, St. Louis, Missouri, USA
[5]AI for Health Institute, Washington University in St. Louis, St. Louis, Missouri, USA

Corresponding Author:
Madelyn R. Frumkin, PhD
Center for Technology and Behavioral Health
Department of Biomedical Data Science
Geisel School of Medicine
Dartmouth College

Email: Madelyn.r.frumkin@dartmouth.edu

Abstract

Ecological Momentary Assessment (EMA) may be useful in the surgical setting for predicting and monitoring symptoms including perioperative pain. However, no largescale studies have established the feasibility of perioperative EMA. The Personalized Prediction of Persistent Postsurgical Pain (P5) study includes collection of perioperative EMA in a cohort of 2,500 patients who underwent major surgery in the Midwestern United States. Despite EMA not being required nor directly incentivized, 91% of participants completed any EMA. Demographic and clinical characteristics were largely similar between those who did and did not complete EMA, except for insurance status. Participants who did not complete any EMA were more likely to be using Medicare (38% vs. 28%) or Medicaid (20% vs. 11%). Without excluding participants based on low compliance, we observed high preoperative (Median = 80%) and postoperative (Median = 72%) compliance with EMAs delivered three times per day. Compliance was somewhat lower among individuals with worse mental and physical health symptoms, though effect sizes were small. Sociodemographic characteristics were associated with both pre- and postoperative EMA compliance, such that men who identified as Black/African heritage responded to fewer than 50% of surveys on average. Individuals who used Medicaid and those with lower educational attainment also exhibited lower compliance. Overall, EMA appears feasible in the surgical setting. However, knowledge of these disparities is critical to ensuring that EMA research is generalizable, as individuals who provide less data will be less influence in statistical models without proper handling of missing data.

Introduction

Ecological Momentary Assessment (EMA) allows for frequent capture of patient-reported symptoms and experiences via smartphone [12,28]. Repeated measurements within daily life reduce recall bias associated with traditional patient-reported outcome measures and improve ecological validity [35,37,39,49]. Despite widespread adoption of EMA in fields including pain medicine, mental health, and substance use [24,27,35,48], relatively few studies have integrated EMA into perioperative assessment. Limited research suggests EMA may be useful for monitoring postoperative recovery and identifying opportunities for early intervention [2,3,13,16,21,34]. Preoperative EMA also facilitates the development of novel features (e.g., symptom trajectories and dynamics) that may improve prediction of surgical outcomes, including persistent post-surgical pain [13,16].

Adoption of remote perioperative monitoring via smartphone is hindered by a lack of largescale studies supporting feasibility. A low response rate may yield insufficient data to accomplish study goals (e.g., dynamic modeling of within-person processes [23]). Lower compliance also increases likelihood that the obtained data points are not representative of the individual's true experience. This occurs when EMA prompts are not missed at random, but due to factors related to the constructs of interests (e.g., if pain reports are more often missed when experiencing increased pain symptoms)[26,33]. Lower compliance from certain demographic groups may also limit generalizability of findings.

Across fields including chronic pain, mental health, and substance use, EMA compliance is consistently estimated in the range of 75-80% [5,9,19,31,47,50–52]. However, 30-50% of EMA studies do not report compliance, and it is common for participants to be excluded from analysis based on unwillingness to complete EMA or low compliance [42,45,50]. Consequently, compliance rates reported in the existing literature may be inflated. Additionally, although existing studies have considered several participant factors including mental health status, age, and gender, no meta-analyses have examined variations in compliance based on other demographic factors, such as education, race, and socioeconomic status. A small number of individual studies suggest compliance may be higher among participants who identify as White [44,46]. If this phenomenon is widespread, then findings derived from EMA data may not generalize to diverse demographic groups without careful handling of missing data.

The goal of the current secondary analysis is to examine demographic and clinical correlates of EMA compliance in a large surgical sample (N = 2,500) that is diverse with respect to factors including race, age, and education. First, we compare demographic and clinical characteristics of individuals who did versus did not provide any amount of EMA. Second, we examine demographic and clinical correlates of EMA compliance, both before and after surgery. Lastly, we examine variability in compliance based on study-related factors, including number of days

before/after surgery, time of day, and assessment schedule. Together, these aims can inform future strategies for maximizing EMA compliance.

Methods

Participants. The Personalized Prediction of Persistent Postsurgical Pain (P5) study is a single-center, prospective study conducted at a tertiary care hospital in the Midwestern United States affiliated with an academic medical center. The overarching goal of P5 is to predict persistent post-surgical pain (PPSP) using multimodal data. Potential participants were identified through the Epic Electronic Health Record (EHR) and screened based on eligibility criteria. Adults aged 18 to 75 undergoing major surgery, defined as surgery duration of >1h, with a planned overnight hospital stay. Access to a smartphone or tablet was required for study participation, although individuals who had access to someone else's device (e.g., spouse) could be included. Individuals unable to speak or read English were excluded. Eligible candidates were contacted via MyChart or direct telephone call before their visit to the hospital's preoperative assessment center. Informed consent was obtained electronically via REDCap or in person. The study was approved by the university's Institutional Review Board. Participants provided written informed consent, and data confidentiality was maintained through secure storage in a REDCap database. Participants were compensated $50 for completion of the preoperative visit and baseline surveys, $25 for completion of the 3-month survey, and $25 for completion of the 6-month survey. Participants were not compensated based on EMA compliance. The P5 is study is approved by the Washington University Institutional Review Board.

Data Collection and Measures. Baseline data were collected prior to surgery and included patient demographics (race, ethnicity, age, sex, education level, and marital status). Participants were asked whether they had ever been diagnosed with anxiety, depression, posttraumatic stress disorder, substance use disorder, chronic pain, and whether they currently used opioid medications for pain management. Participants were also asked to rate their pain at rest and during movement in the past week on a scale from 0 (no pain) to 10 (worst possible pain). Participants completed validated patient-reported outcome measures, including the Patient-Reported Outcomes Measurement Information System (PROMIS®) short-form surveys for anxiety, depression, physical functioning, sleep disturbance, and cognitive abilities [6], as well as the Pain Catastrophizing Scale (PCS) [30,41]. Cognitive function was further assessed using the Color-Word Stroop Test (CWST) [40]. Insurance status and postoperative pain intensity were extracted from the EHR. Postoperative pain intensity was defined as the median of all pain ratings collected in the Post Anesthesia Care Unit (PACU), typically within the first 3 hours after surgery.

EMA data were collected using the LifeData app downloaded onto the participant's personal smartphone. Participants received EMA prompts three times daily. Prompts were delivered 5 hours apart. Participants chose to begin receiving EMA prompts at 7am, 8am, or 9am. Each

EMA contained 15 items assessing pain intensity, interference, momentary catastrophizing, fatigue, depressed mood, anxiety, and pain medication craving and use (see **Supplementary Table S1**). The first survey of the day contained an additional item assessing sleep quality. All items had a 0-100 visual analoge response scale. EMA typically began on the day of enrollment and continued through 30 days postoperatively. Participants were instructed to delete the EMA app on day 30. Participants could self-withdraw from EMA before day 30 by deleting the smartphone application. Participants were not withdrawn from the overall study if they stopped receiving and/or completing EMAs.

Statistical Analysis. First, we compared demographic and clinical characteristics of individuals who did versus did not provide any amount of EMA (1+ survey) using chi-square tests for categorical predictors and nonparametric Wilcoxon rank sum for continuous predictors.

Second, we characterized pre- and postoperative EMA compliance among participants who received any EMA prompts (including individuals who did not complete any prompts). Compliance was calculated for each individual as the number of completed EMAs divided by the total number of possible EMAs in the pre- and postoperative periods. Preoperatively, possible EMAs was defined as the number of days enrolled prior to surgery times three (for three daily observations). Because some participants had longer preoperative windows due to extenuating circumstances (e.g., delayed surgery), we limited the preoperative window to two weeks before surgery. Postoperatively, all participants had 87 possible EMAs (29 days of three daily observations).

For pre- and postoperative compliance, we report the mean, median, interquartile range (IQR), and range across the sample. We also examined demographic and clinical correlates of compliance using nonparametric Wilcoxon rank sum for two-level categorical predictors (e.g., current opioid use), Kruskal-Wallis H test for categorical predictors with more than two levels (e.g., type of insurance), and Spearman correlations for continuous predictors (e.g., age). Bonferroni correction was used in all analyses to reduce risk of false positives. Effect sizes are reported as r for Wilcoxon rank sum and Spearman correlations, where 0.1, 0.3, and 0.5 indicate small, medium, and large effects, respectively. Effect sizes are reported as η^2 for Kruskal-Wallis, where 0.01, 0.06, and 0.14 indicate small, medium, and large effects, respectively[43].

In addition to examining bivariate relationships between predictors and compliance, we used random forest models to explore combined predictive value, as well as possible interactions among predictors (e.g., race and preoperative mental health diagnoses). Given small cell sizes in some categories, only participants who identified their race as Black/African heritage or White/Caucasian were included. Missing predictor values were imputed using the missForest algorithm, a nonparametric imputation method that iteratively predicts missing values via random forest models. The imputed dataset was used to train a random forest regression model using the rfsrc() function from the randomForestSRC package in R with default parameters (500

trees, √p variables considered at each split). Separate models were constructed for predicting pre- and postoperative compliance. Model performance was estimated using out-of-bag (OOB) error and pseudo R^2. To explore potential interactions among predictors, pairwise interaction strengths were extracted using the find.interaction() function. The resulting interaction matrix was visualized using a heatmap. The strongest interactions were investigated via linear regression to understand directionality.

Finally, we examined variability in compliance based on study-related characteristics. Logistic regression was used to examine whether likelihood of response differed by day relative to surgery, where negative numbers are number of days before surgery (e.g., -7 = 7 days before surgery) and positive number are number of days after surgery (e.g., 7 = 7 days after surgery). Logistic regression was also used to examine whether likelihood of response differed by time of day based on morning, early afternoon, and early evening survey groups. Finally, linear regression models were also used to examine whether compliance differed based on assessment schedule. Here, participants were grouped into 3 categories: those who chose to initiate surveys at 7am, 8am, or 9am. All analyses were performed in R version 4.4.1.

Results

A total of 2500 participants signed consent to enroll in the parent study. Of these, 107 participants (4.3%) become ineligible prior to surgery for reasons including canceled or changed surgery and non-completion of baseline measures. Another 64 participants (2.6%) asked to withdraw prior to surgery. The remaining 2329 participants were majority female (n = 1555, 66.8%) and identified as White/Caucasian (n = 1752, 75.2%) or Black/African heritage (n = 427, 18.3%). Participants had an average age of 53 years (SD = 13.8, Min = 18, Max = 75).

Most participants (n = 2127, 91%) completed some pre- or postoperative EMA. As shown in **Table 1**, demographic and clinical characteristics were largely similar between those who did and did not provide EMA responses. After correction for multiple tests, the only significant difference between the two groups was insurance status. Participants who did not complete any EMA were more likely to be using governmentally-subsidized insurance, such as Medicare (38% of individuals without EMA, vs. 28% of those with EMA) or Medicaid (20% of individuals without EMA, vs. 11% of those with EMA; $\chi^2(4) = 27.32$, $p < .001$). Participants who did not complete any EMA appeared somewhat older (Median age = 54 vs. 58) and reported higher past week pain at rest (Median = 3.0 vs. 4.5) and pain with activity (Median = 5.0 vs. 6.0). However, these effects were not significant after correction for multiple tests. The two groups did not differ with regard to surgical site, sex, race, ethnicity, education, marital status, or severity of other physical, psychosocial, or cognitive symptoms.

Preoperative compliance. A total of 2171 participants received at least 1 EMA prior to surgery. Participants received an average of 28 surveys (SD = 14) and completed an average of 19.5

surveys (SD = 14) in the two weeks before surgery. The average preoperative compliance rate was 68% (SD = 32%, Median = 80%). As shown in **Figure 1**, the distribution of preoperative compliance was left-skewed (skewness = -0.95).

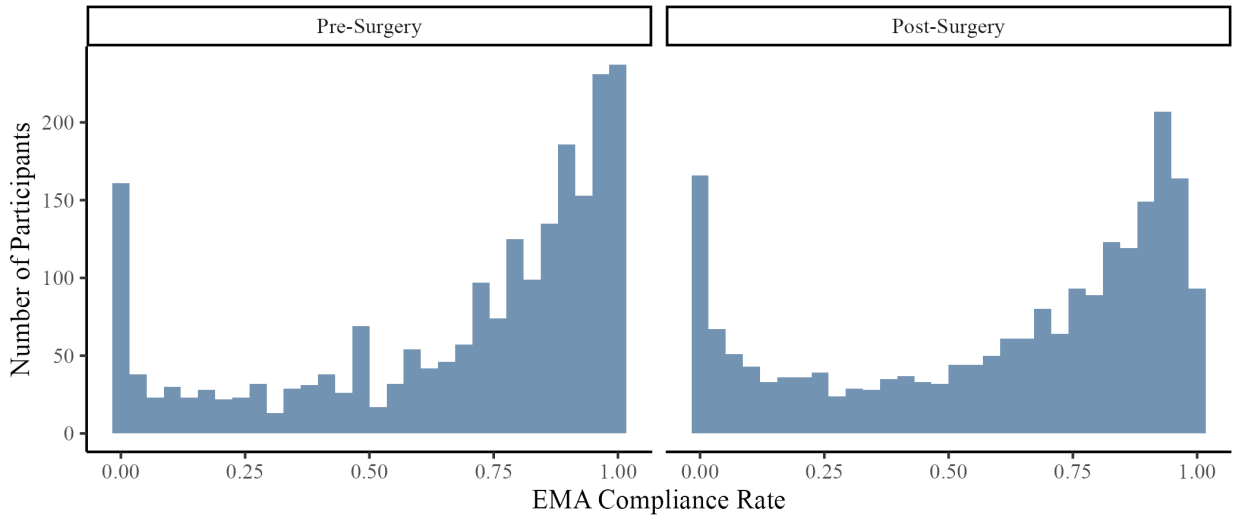

Figure 1. Histograms of pre- and postoperative compliance rates. EMA = ecological momentary assessment.

Correlations between preoperative compliance and demographic and clinical characteristics are visualized in **Fig 2a**. Preoperative EMA compliance was somewhat lower among participants reporting worse preoperative pain catastrophizing (r = -.11, p.adj < .001), PROMIS anxiety (r = -.09, p.adj < .001), PROMIS depression (r = -.08, p = .002), pain at rest on the numeric rating scale (r = -.07, p.adj = .008), and subjective cognitive function based on the PROMIS (r = .07, p.adj = .014).

Group differences are visualized in **Fig 2b**. Significant differences in preoperative EMA compliance were observed based on race (η^2 = .028, p.adj < .001), education (η^2 = .016, p.adj < .001), type of insurance (η^2 = .015, p.adj < .001), and marital status (η^2 = .011, p.adj = .036). Compliance was higher among White/Caucasian (n = 1639, M = 70%, Median = 82.5%, IQR = 54%-94%) than Black participants (n = 394, M = 58%, Median = 69%, IQR = 28%-88%). Below average compliance was also observed among participants with less than a high school education (n = 116, M = 58%, Median = 69%, IQR = 36%-88%), individuals who reported their marital status as separated (n = 45, M = 59%, Median = 61%, IQR = 39%-86%) or domestic partnership (n = 47, M = 57%, Median = 61%, IQR = 33%-87%), and those using Medicaid (n = 246, M = 57%, Median = 69%, IQR = 26%-86%). Preoperative compliance did not appear to differ by sex, surgery site, current use of opioids, or based on mental health, substance use, or chronic pain diagnoses (see **Supplementary Table S2**).

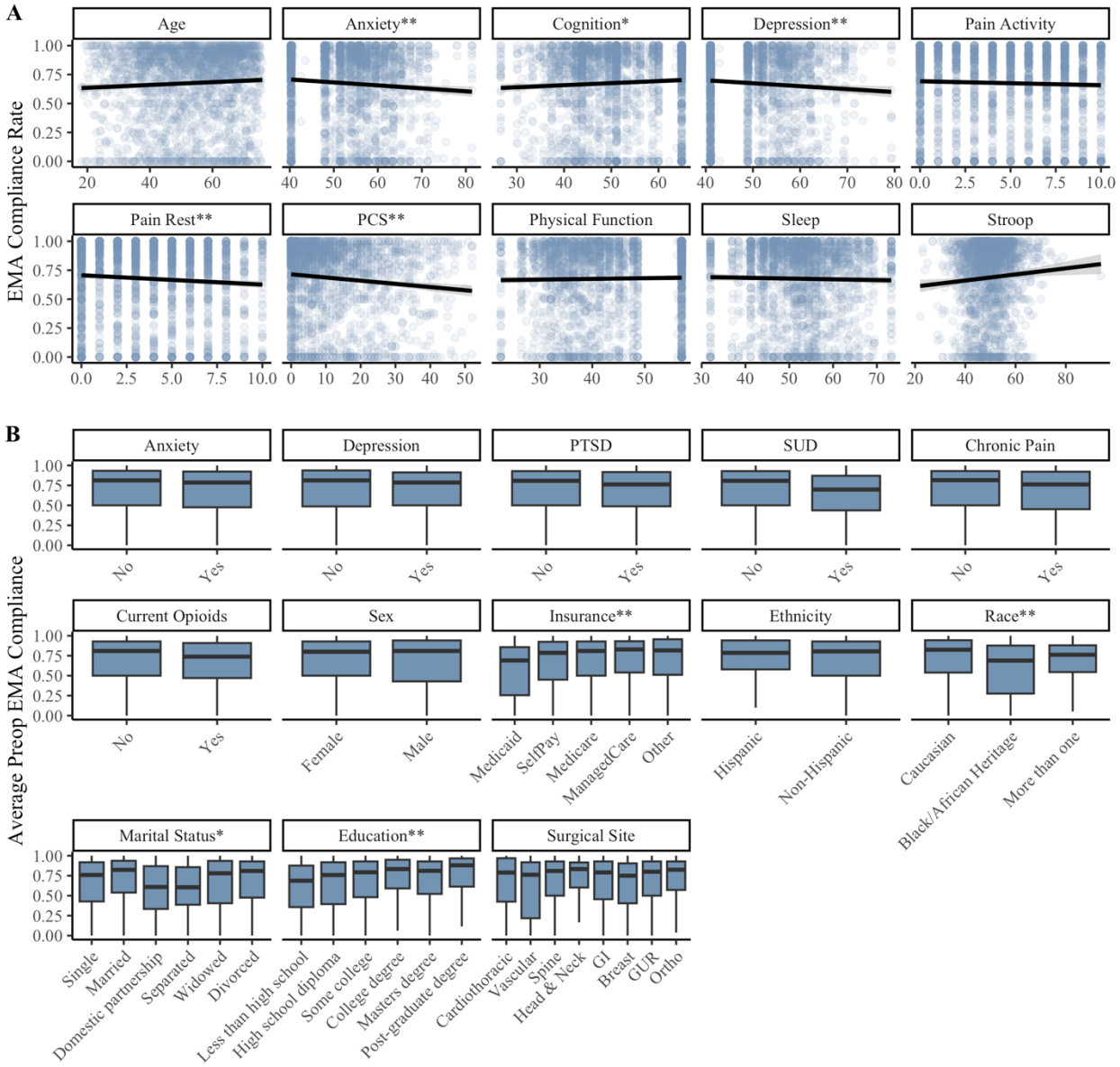

Figure 2. Associations between preoperative Ecological Momentary Assessment (EMA) compliance and demographic and clinical characteristics. Panel **A** shows correlations between preoperative compliance and continuous predictors. Panel **B** shows box plots of preoperative compliance by categorical predictors (midline = median). **Bonferroni-corrected $p < .01$; *Bonferroni-corrected $p < .05$; PCS = Pain Catastrophizing Scale; PTSD = Post-Traumatic Stress Disorder; SUD = Substance Use Disorder; GI = gastrointestinal; GUR = Genitourinary. In Panel A, anxiety, subjective cognition, depression, physical function, and subjective sleep impairment were assessed via PROMIS self-report measures; pain with activity and pain at rest were assessed via numeric rating scales. In Panel B, anxiety, depression, PTSD, SUD, and chronic pain refer to self-reported lifetime diagnoses.

Postoperative compliance. Among the 2130 participants who received at least 1 postoperative EMA, the average postoperative compliance was 61% (SD = 33%, Median = 72%). Participants received an average of 81 surveys (SD = 17) and completed an average of 53 surveys (SD = 29) in the 29 days following surgery. As shown in **Figure 1b**, the distribution of compliance rates was moderately left-skewed (skewness = -0.65).

Correlations between postoperative compliance and demographic and clinical characteristics are visualized in **Fig 3a**. Postoperative EMA compliance was lower among participants reporting worse preoperative pain catastrophizing ($r = -.17$, p.adj < .001), depression ($r = -.17$, p.adj < .001), anxiety ($r = -.15$, p.adj < .001), pain at rest ($r = -.13$, p.adj < .001), subjective cognitive function ($r = .11$, p.adj < .001), activity-related pain ($r = -.10$, p.adj < .001), subjective sleep impairments ($r = -.08$, p.adj = .002), and physical function ($r = .07$, p.adj = .022). Postoperative EMA compliance was also lower among participants who reported greater acute postoperative pain intensity in the PACU ($r = -.15$, p.adj < .001). Postoperative EMA compliance was higher among older participants ($r = .11$, p.adj < .001). All effects were small.

Group differences are visualized in **Fig 3b**. Consistent with preoperative findings, significant differences in postoperative EMA compliance were observed based on race ($\eta^2 = .021$, p.adj < .01), education ($\eta^2 = .034$, p.adj < .01), type of insurance ($\eta^2 = .015$, p.adj < .01), and marital status ($\eta^2 = .019$, p.adj < .01). Additional group differences were observed based on surgical site ($\eta^2 = .028$, p.adj < .01), with below average postoperative compliance observed among patients having vascular (n = 31, M = 51%, Median = 62%, IQR = 19-79%) and cardiothoracic (n = 224, M = 52%, Median = 64%, IQR = 14-84%) procedures. A lifetime history of depression ($\eta^2 = .015$, p.adj < .01), anxiety ($\eta^2 = .009$, p.adj < .01), chronic pain ($\eta^2 = .009$, p.adj < .01), and substance use disorders ($\eta^2 = .005$ p.adj = .047) were also associated with lower postoperative EMA compliance. Additionally, participants who reported current opioid use prior to surgery had lower postoperative compliance ($\eta^2 = .006$, p.adj = .017). Effect sizes remained small (e.g., $\eta^2 < .06$; see **Supplementary Table S3**).

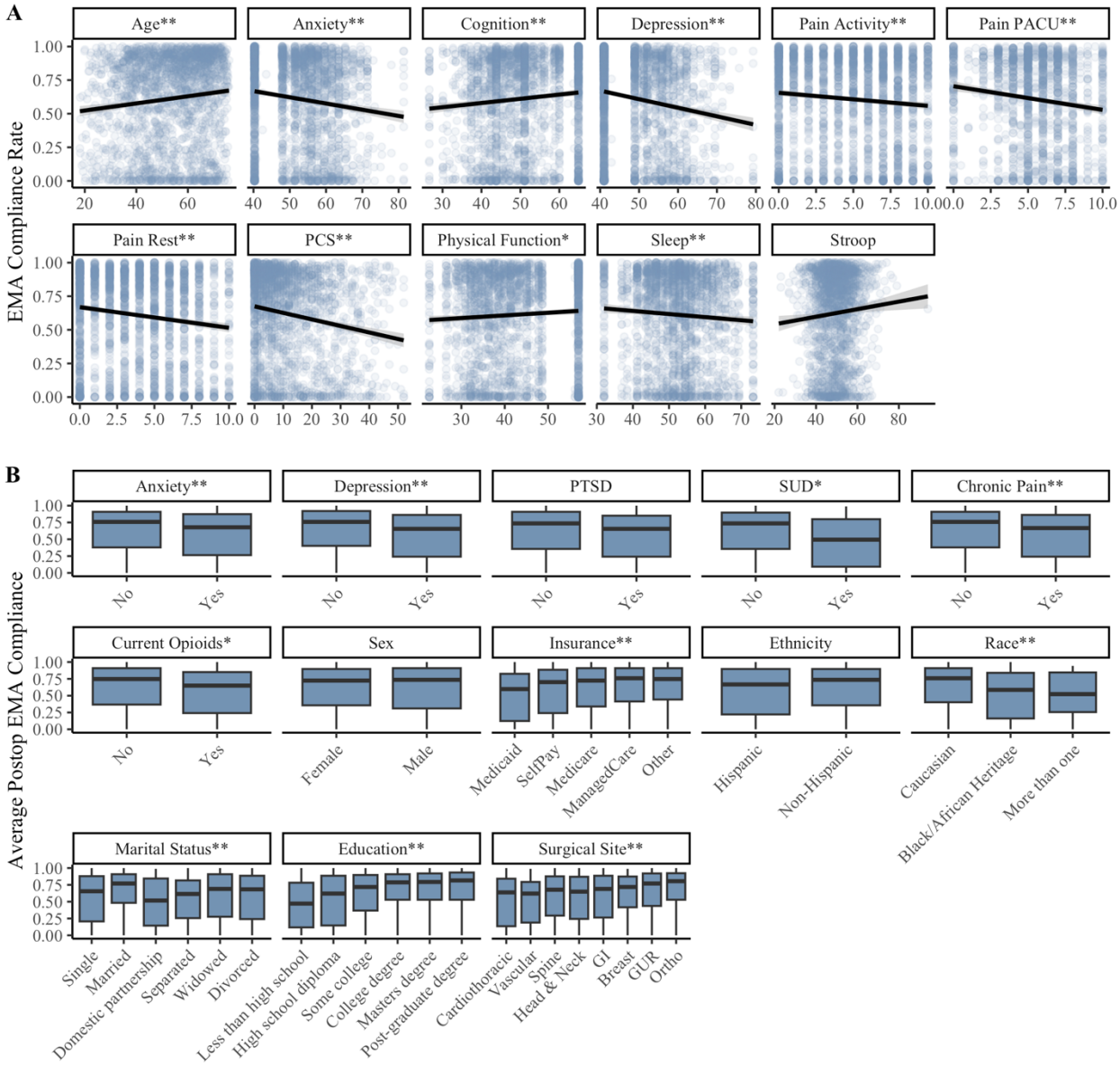

Figure 3. Associations between postoperative Ecological Momentary Assessment (EMA) compliance and demographic and clinical characteristics. Panel **A** shows correlations between postoperative compliance and continuous predictors. Panel **B** shows box plots of postoperative compliance by categorical predictors (midline = median). **Bonferroni-corrected $p < .01$; *Bonferroni-corrected $p < .05$; PACU = Post Anesthesia Care Unit; PCS = Pain Catastrophizing Scale; PTSD = Post-Traumatic Stress Disorder; SUD = Substance Use Disorder; GI = gastrointestinal; GUR = Genitourinary. In Panel A, anxiety, subjective cognition, depression, physical function, and subjective sleep impairment were assessed via PROMIS self-report measures; pain with activity and pain at rest were assessed via numeric rating scales. In Panel B, anxiety, depression, PTSD, SUD, and chronic pain refer to self-reported lifetime diagnoses.

Combined influence of predictors on EMA compliance. For preoperative EMA compliance, the random forest regression accounted for approximately 1.6% of the variance in compliance rates (OOB error = 10.1%). For postoperative EMA compliance, the random forest regression accounting for approximately 9.7% of the variance in compliance rates (OOB error = 10.0%). Both pre- and postoperatively, demographic features (e.g., race and sex) appeared to interact with one another and with lifetime diagnoses of mental health, chronic pain, and substance use disorders (see **Figure 4A-B**). In follow-up linear regression analyses, the only significant interaction was between sex and race, with Black/African heritage men (n = 95) exhibiting low EMA compliance both preoperatively (M = 47.6%, Median = 47.5%, IQR = 9-81%) and postoperatively (M = 43.0%, Median = 34.5%, IQR = 4-80%; see **Figure 4C**).

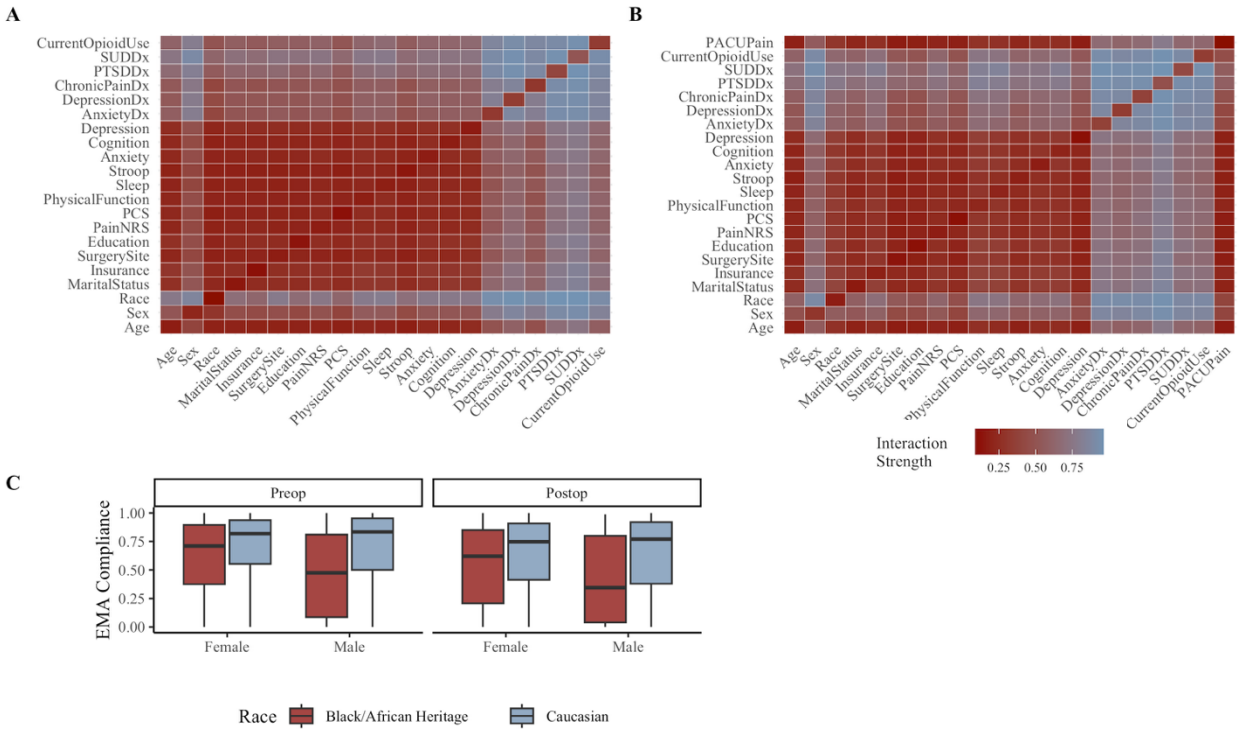

Figure 4. Interaction effects among predictors of Ecological Momentary Assessment (EMA) compliance. Panel **A** is a heatmap of interaction strength derived from a random forest regression model predicting preoperative EMA compliance. Panel **B** is a heatmap of interaction strength derived from a random forest regression model predicting postoperative EMA compliance. Panel **C** displays the only statistically significant interaction (race x sex) in follow-up linear regression analysis. PACU = Post Anesthesia Care Unit; PCS = Pain Catastrophizing Scale; PTSD = Post-Traumatic Stress Disorder; SUD = Substance Use Disorder; NRS = numeric rating scale (average of pain at rest and pain with activity). Predictors labeled 'Dx' refer to lifetime diagnoses. Predictors labeled 'Depression,' 'Cognition,' 'Anxiety,' 'Sleep,' and 'PhysicalFunction' are PROMIS measures.

Influence of EMA schedule on compliance. Most participants (65%) opted to begin receiving EMAs at 9am. However, 15.5% of participants opted to begin assessments at 7am, and 19.5% opted to begin assessments at 8am. Assessment schedule was a significant predictor of pre- and postoperative compliance (**see Figure 5A**). Compared to those whose EMAs began at 9am, participants who opted to receive assessments beginning at 7am or 8am had increased preoperative compliance by approximately 4-5% ($p = .004$ and .031, respectively; $R^2 = 0.005$). Similarly, participants who opted to receive assessments beginning at 7am or 8am had increased postoperative compliance by approximately 6-9% ($p < .001$ and .002, respectively; $R^2 = 0.015$). Postoperatively, participants with assessments beginning at 8am were most compliant.

Influence of time-related variables on compliance. Preoperatively, compliance with evening surveys was more likely than compliance with morning surveys ($\beta = 0.095$, SE = 0.02, $z = 4.44$, $p < .001$, OR = 1.10). Postoperatively, compliance with afternoon surveys was more likely than compliance with morning surveys ($\beta = 0.135$, SE = 0.01, $z = 11.00$, $p < .001$, OR = 1.15), and compliance with evening surveys was also more likely than compliance with morning surveys ($\beta = 0.201$, SE = 0.01, $z = 16.25$, $p < .001$, OR = 1.22).

As shown in **Figure 5B**, compliance was relatively stable in the preoperative period. There was no effect of day relative to surgery on likelihood of completing EMA ($\beta = -0.003$, SE = .002, $p = .187$). In the postoperative period, there was a linear effect of day since surgery ($\beta = 0.02$, SE = 0.003, $z = 8.74$, $p < .001$), suggesting that odds of responding increased by approximately 2% per day after surgery (OR = 1.02). However, the quadratic term was also significant ($\beta = -0.0006$, SE = 0.000081, $z = -7.41$, $p < .001$; OR = 0.99), indicating a nonlinear inverted U-shaped relationship, whereby compliance initially increased (e.g., days 1-3 in Figure 5B) then later decreased somewhat across the postoperative period. Additionally, 288 participants (13.5%) self-unenrolled from EMA in the postoperative period (red line; Figure 5B).

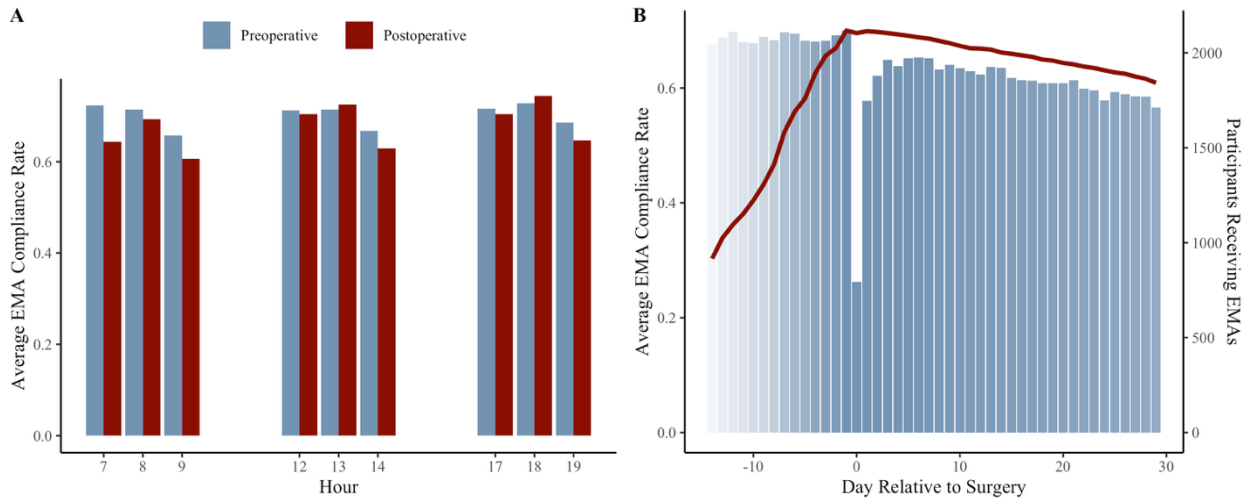

Figure 5. Time-related predictors of Ecological Momentary Assessment (EMA) compliance. Panel **A** compares average compliance rates across assessment schedules. Most participants (65%) retained the default assessment schedule beginning at 9am. However, 15.5% and 19.5% opted to begin assessments at 7am and 8am, respectively. Panel **B** shows average EMA compliance by day relative to surgery (lefthand Y-axis). Blue bars capture compliance calculated as surveys completed / surveys received. The red line indicates number of participants receiving EMAs on a given day relative to surgery (righthand Y-axis). The redline decreases in the postoperative period because some participants (n = 288; 13.5%) self-unenrolled from EMA by deleting the app.

Discussion

Digital technologies are increasingly used for remote monitoring of surgical patients [11,20]. Non-invasive wearable devices provide continuous, objective data relevant to physical activity, cardiovascular functioning, and sleep [22]. Wearable devices have been applied especially in the postoperative setting to monitor physical activity for early signs of impaired recovery [1,20]. Other key surgical outcomes, including persistent pain, subjective impairment, mental and cognitive health symptoms, and fatigue cannot be directly passively monitored [7]. Passive sensors contained on wearable devices and smartphones can be used to *predict* fluctuations in symptoms, especially mood and anxiety [4]. However, there are currently no validated objective markers of pain, a key postoperative domain associated with increased morbidity, prolonged opioid use, impaired functional recovery, and higher healthcare costs [8,14,15,36]. Thus, remote perioperative monitoring via smartphone can lead to improved understanding, prediction, and prevention of persistent postoperative pain and impairment.

This study establishes the feasibility of perioperative EMA in a large surgical cohort. Without removing individuals based on low compliance, average preoperative compliance with EMA delivered three times daily was 68% (Median = 80%), and average postoperative compliance was 61% (Median = 72%) in the first 30 days after surgery. Average compliance rates were lower than the estimated 75-80% found in numerous meta-analyses across general and specific populations [5,9,19,31,47,50–52]. This study is the first to assess EMA compliance in a largescale surgical cohort. Our findings could suggest lower compliance in the perioperative setting, compared to other samples. Indeed, it is easy to see how completing EMA could be challenging perioperatively. However, unlike the majority of studies included in prior meta-analyses, we did not exclude participants based on low compliance. Instead, we included any individual who *received* at least one EMA prompt, even those who completed no prompts. If we were to remove participants based no or low compliance (e.g., <20% compliance), our compliance would be consistent with prior meta-analyses. Especially given that many EMA studies with similar compliance rates involve healthy undergraduates and pay participants per survey, this level of compliance is impressive.

Compliance with EMA is a concern given that individuals unwilling or unable to complete EMA will be unrepresented in statistical models, and those who provide fewer data points will be less influential. If compliance varies systematically by demographic or clinical characteristics, this may result in less generalizable findings. Given that participants were not excluded from the larger study based on EMA noncompliance, we were able to examine demographic and clinical characteristics of those who did versus did not provide EMA. After correcting for multiple tests, the only significant difference between the two groups was insurance status. Participants who did not complete any EMA were more likely to be using Medicare or Medicaid. Though not a perfect proxy for socioeconomic status, adults under the age of 65 are eligible for Medicaid with annual household income below approximately $21,000 for single individuals, and $43,000 for a family of four. Future studies should investigate barriers to participating in EMA for lower income individuals.

Individuals using Medicaid also exhibited lower pre- and postoperative compliance, as did participants who identified as Black/African heritage, and those with lower educational attainment. In probing possible interaction effects among demographic and clinical characteristics, we found Black/African heritage men to have the lowest EMA compliance, completing fewer than 50% of surveys on average. To our knowledge, no meta-analyses have demonstrated sociodemographic differences in EMA compliance. However, some individual studies have demonstrated this pattern in small samples [44,46]. The demographic discrepancies we found may have implications for the generalizability of EMA, especially given that missing EMA observations are often not imputed [12]. Notably, there was large within-group variability, resulting in small effect sizes.

Preoperative EMA compliance did not differ by sex, surgery site, current use of opioids, or based on lifetime history of any mental health, substance use, or chronic pain diagnoses. We found small effects of several mental and physical health symptoms, such that individuals with worse symptoms tended to have slightly lower EMA compliance. These findings are consistent with prior meta-analyses suggesting compliance does not appear to be lower among healthy versus clinical samples, though there may be within-sample effects of emotional and cognitive impairments [19,50,52]. However, the effects we observed were small in magnitude, explaining little variance in preoperative EMA compliance.

Postoperative EMA compliance did differ by surgical site. Patients who underwent vascular and cardiothoracic procedures exhibited lower postoperative compliance, whereas patients who underwent breast, genitourinary, or orthopedic surgery demonstrated the highest postoperative compliance. This difference may be due in part to variation in acute postoperative impairment, as acute pain intensity recorded in the post-anesthesia care unit was also weakly associated with postoperative EMA compliance. Our analyses did not account for times when the participant may have been unable to respond (e.g., due to prolonged sedation).

A lifetime history of several mental health, substance use, and chronic pain diagnoses were also associated with lower postoperative EMA compliance. Additionally, participants who reported preoperative opioid use had lower postoperative compliance. Individuals with worse mental and physical health symptoms again exhibited somewhat lower postoperative EMA compliance. Effect sizes remained small, though correlations were somewhat larger than those assessed preoperatively. Consistent with prior meta-analyses [29,51,52], we found postoperative EMA compliance to be positively associated with age, such that older adults exhibited somewhat *higher* compliance. In the random forest model, combined predictors accounted for approximately 10% of the variance in postoperative EMA compliance.

In addition to participant characteristics, we also examined the impact of EMA schedule on compliance. Participants who opted for earlier start times tended to exhibit higher compliance. Most advice on improving compliance with EMA is anecdotal, rather than empirical [12]. It is possible that starting EMA earlier could improve compliance. However, it is equally possible that participants who opted for earlier start times possessed other characteristics associated with EMA compliance, such as higher conscientiousness or better overall health. Compliance with evening surveys tended to be most likely. Thus, if choosing sparser assessment schedules (e.g., daily diary), evening assessment may be preferred to morning assessment.

Finally, we observed consistent compliance across the days leading up to surgery. After surgery, there was an initial increase in compliance over the first few days when many patients would have still been hospitalized and recovering from the acute effects of surgery. Postoperative compliance began to decline after about one week, and approximately 13.5% of participants self-

unenrolled from EMA before 30 days by deleting the application. Notably, participants in our study were not paid based on EMA compliance, which some studies suggest improves sustained compliance [10,19,47,50,51].

This study is the first to our knowledge to comprehensively assess feasibility of EMA in the perioperative setting. We recruited a large sample that was diverse with respect to race, educational attainment, and socioeconomic status—factors which have not previously been examined as predictors of EMA compliance in meta-analyses due to homogeneity and small sample sizes of individual studies [5,9,19,31,47,50–52]. Participants were not paid directly for EMA or excluded from analyses based on low or no EMA completion. Thus, we view our compliance estimates as highly likely to translate to real-world applications of EMA outside of the research context.

Our study also has limitations. Given that all participants underwent surgery, findings regarding demographic and clinical correlates of compliance may not replicate to all settings. Findings may also be specific to our study procedures, including the EMA items, length of the assessments, assessment schedule, and length of the study, though there is limited evidence that these study-related factors impact compliance [10,19,47,50,51]. Although our sample was larger and more diverse than prior EMA studies [5,9,19,31,47,50–52], there was a lack of representation of some identities, including races other than White/Caucasian or Black/African heritage. Based on 2020 census data, St. Louis City and the surrounding St. Louis County are approximately 4.6% Asian, 28.8% Black/African heritage, and 4% Hispanic or Latino [38]. Thus, although we anticipated a higher proportion of White/Caucasian and Black/African heritage participants, non-White and Hispanic/Latino individuals remain underrepresented compared to local population statistics. We also did not collect information about some demographic characteristics, including income and work status, which may be informative for characterizing EMA compliance. Finally, although the vast majority of Americans own a smartphone [32], this inclusion criteria for the parent study may limit generalizability.

Overall, remote perioperative monitoring via smartphone appears feasible. Demographic characteristics including race, insurance status, and education were robustly associated with perioperative EMA compliance. This phenomenon has not previously been investigated in any largescale studies or meta-analyses of EMA compliance. Without appropriate handling of missing data, models may be systematically biased towards the experiences on individuals from groups who are historically over-represented in research, liming generalizability to individuals who hold marginalized identities. Despite availability of techniques including missing data imputation, Full Information Maximum Likelihood (FIML) estimation, and Kalman filtering, these techniques are not widely used in EMA research [12,17,18,25]. Given rapid adoption of remote monitoring via EMA, future research should formalize user-friendly strategies for

handling missing data and employ community-engaged research practices to understand barriers and facilitators of remote monitoring via smartphone.

Table 1. Demographic and clinical characteristics of participants who did and did not complete any EMA

Predictor	Levels	Has EMA *N = 2127 (91%)*	No EMA *N = 202 (9%)*	p	p.adj
Race	American Indian/ Alaskan Native	15 (1%)	1 (0%)	0.515	1
	Asian	14 (1%)	0 (0%)		
	Black/African Heritage	383 (18%)	44 (22%)		
	Caucasian	1608 (76%)	144 (72%)		
	Hawaiian Native/Other Pacific Islander	4 (0%)	0 (0%)		
	More than one	33 (2%)	5 (2%)		
	Other	30 (1%)	2 (1%)		
	Prefer not-answer	23 (1%)	4 (2%)		
Ethnicity	Hispanic or Latino	47 (2%)	4 (2%)	1	1
	Non-Hispanic or Latino	1984 (98%)	186 (98%)		
Age	Median (IQR)	54.35 (42.79-64.50)	58.23 (44.57-67.16)	0.008	0.176
Sex	Female	1430 (67%)	125 (62%)	0.143	1
	Male	697 (33%)	77 (38%)		
Education	College degree	455 (22%)	35 (18%)	0.123	1
	High school diploma	480 (23%)	52 (26%)		
	Less than high school	113 (5%)	17 (9%)		
	Master's degree	299 (14%)	19 (10%)		
	Post-graduate degree	90 (4%)	10 (5%)		
	Some college	679 (32%)	66 (33%)		
Marital Status	Divorced	239 (11%)	31 (16%)	0.327	1
	Domestic partnership	44 (2%)	7 (4%)		
	Married	1152 (55%)	97 (49%)		
	Other	8 (0%)			
	Single	512 (24%)	51 (26%)		
	Widowed	108 (5%)	10 (5%)		
	separate	46 (2%)	3 (2%)		
Anxiety	Yes	867 (41%)	87 (44%)	0.533	1
Depression	Yes	795 (38%)	89 (45%)	0.055	1
PTSD	Yes	229 (11%)	23 (12%)	0.841	1
SUD	Yes	72 (3%)	8 (4%)	0.814	1
Chronic Pain	Yes	635 (30%)	70 (36%)	0.142	1
Current Opioid Use	Yes	280 (14%)	37 (19%)	0.049	1
Pain at Rest	Median (IQR%)	3.00 (1.00-6.00)	4.50 (1.00-7.00)	0.005	0.115
Pain with Activity	Median (IQR%)	5.00 (2.00-7.00)	6.00 (2.75-8.00)	0.008	0.18
PCS	Median (IQR%)	8.00 (2.00-18.00)	10.00 (4.00-23.00)	0.018	0.423
PROMIS Cognition	Median (IQR%)	49.60 (43.80-58.80)	50.90 (43.62-56.18)	0.479	1
PROMIS PF	Median (IQR%)	41.00 (34.70-48.60)	38.65 (34.12-45.18)	0.014	0.322
PROMIS Anxiety	Median (IQR%)	52.50 (40.30-59.48)	54.10 (40.30-60.30)	0.113	1
PROMIS Depression	Median (IQR%)	48.90 (41.00-55.90)	49.55 (41.00-58.48)	0.029	0.675
PROMIS Sleep	Median (IQR%)	52.80 (46.40-58.08)	52.80 (48.30-56.53)	0.578	1
Stroop	Median (IQR%)	49.00 (45.00-54.00)	48.00 (45.00-52.00)	0.042	0.955
Surgical Site	Breast	57 (3%)	4 (2%)	0.449	1
	Cardiothoracic	222 (10%)	23 (11%)		
	Gastrointestinal	448 (21%)	42 (21%)		
	Genitourinary	621 (29%)	52 (26%)		

	Head & Neck	123 (6%)	6 (3%)		
	Ortho	377 (18%)	44 (22%)		
	Other	21 (1%)	4 (2%)		
	Spine	229 (11%)	23 (11%)		
	Vascular	29 (1%)	4 (2%)		
Insurance	Managed Care	978 (46%)	66 (33%)	<.001	<.001
	Medicaid	230 (11%)	39 (20%)		
	Medicare	599 (28%)	75 (38%)		
	Other	96 (5%)	5 (2%)		
	Self-Pay	218 (10%)	15 (8%)		

Note. p.adj = Bonferroni-corrected p-value; PTSD = Post-Traumatic Stress Disorder; SUD = Substance Use Disorder; PCS = Pain Catastrophizing Scale; PROMIS = Patient-Reported Outcomes Measurement Information System; PF = Physical Function; Stroop = Color-Word Stroop Test. Anxiety, depression, PTSD, SUD, and chronic pain refer to self-reported lifetime diagnoses. Past week pain with activity and pain at rest were assessed via 0-10 numeric rating scales.

Acknowledgements

The study was supported by a Congressionally Directed Medical Research Programs grant from the US Department of Defense (W81XWH-21-1-073).

SH has received personal fees from Vertex, unrelated to the current study. TLR receives funding from the National Institutes of Health and American Cancer Society, unrelated to the current study. Other authors report no conflicts of interest.

References

[1] Amin T, Mobbs RJ, Mostafa N, Sy LW, Choy WJ. Wearable devices for patient monitoring in the early postoperative period: a literature review. Mhealth 2021;7:50.

[2] Aminpour E, Holzer KJ, Frumkin M, Rodebaugh TL, Jones C, Haroutounian S, Fritz BA. Preoperative predictors of acute postoperative anxiety and depression using ecological momentary assessments: a secondary analysis of a single-centre prospective observational study. British Journal of Anaesthesia 2025;134:102–110.

[3] Bond DS, Thomas JG, Jones DB, Schumacher LM, Webster J, Evans EW, Goldschmidt AB, Vithiananthan S. Ecological momentary assessment of gastrointestinal symptoms and risky eating behaviors in Roux-en-Y gastric bypass and sleeve gastrectomy patients. Surgery for Obesity and Related Diseases 2021;17:475–483.

[4] Bryan AC, Heinz MV, Salzhauer AJ, Price GD, Tlachac ML, Jacobson NC. Behind the Screen: A Narrative Review on the Translational Capacity of Passive Sensing for Mental Health Assessment. Biomedical Materials & Devices 2024;2:778–810.

[5] Cain AE, Depp CA, Jeste DV. Ecological momentary assessment in aging research: A critical review. Journal of Psychiatric Research 2009;43:987–996.

[6] Cella D, Choi SW, Condon DM, Schalet B, Hays RD, Rothrock NE, Yount S, Cook KF, Gershon RC, Amtmann D. PROMIS® adult health profiles: efficient short-form measures of seven health domains. Value in health 2019;22:537–544.

[7] Chow A, Mayer EK, Darzi AW, Athanasiou T. Patient-reported outcome measures: The importance of patient satisfaction in surgery. Surgery 2009;146:435–443.

[8] Cowen R, Stasiowska MK, Laycock H, Bantel C. Assessing pain objectively: the use of physiological markers. Anaesthesia 2015;70:828–847.

[9] Davanzo A, d´Huart D, Seker S, Moessner M, Zimmermann R, Schmeck K, Behn A. Study Features and Response Compliance in Ecological Momentary Assessment Research in Borderline Personality Disorder: Systematic Review and Meta-analysis. Journal of Medical Internet Research 2023;25:e44853.

[10] Edney S, Goh CM, Chua XH, Low A, Chia J, Koek DS, Cheong K, Dam R van, Tan CS, Müller-Riemenschneider F. Evaluating the Effects of Rewards and Schedule Length on Response Rates to Ecological Momentary Assessment Surveys: Randomized Controlled Trials. Journal of Medical Internet Research 2023;25:e45764.

[11] Farias FAC de, Dagostini CM, Bicca Y de A, Falavigna VF, Falavigna A. Remote Patient Monitoring: A Systematic Review. Telemedicine and e-Health 2020;26:576–583.

[12] Fritz J, Piccirillo ML, Cohen ZD, Frumkin M, Kirtley O, Moeller J, Neubauer AB, Norris LA, Schuurman NK, Snippe E, Bringmann LF. So You Want to Do ESM? 10 Essential Topics for Implementing the Experience-Sampling Method. Advances in Methods and Practices in Psychological Science 2024;7:25152459241267912.

[13] Frumkin MR, Greenberg JK, Boyd P, Javeed S, Shayo B, Shin J, Wilson EA, Zhang JK, Sullivan MJL, Haroutounian S, Rodebaugh TL. Establishing the reliability, validity, and prognostic utility of the Momentary Pain Catastrophizing Scale for use in ecological momentary assessment research. The Journal of Pain 2023;28:1423–1433.

[14] Gan TJ. Poorly controlled postoperative pain: prevalence, consequences, and prevention. Journal of Pain Research 2017;10:2287–2298.

[15] Gan TJ, Habib ,Ashraf S., Miller ,Timothy E., White ,William, and Apfelbaum JL. Incidence, patient satisfaction, and perceptions of post-surgical pain: results from a US national survey. Current Medical Research and Opinion 2014;30:149–160.

[16] Greenberg JK, Frumkin M, Xu Z, Zhang J, Javeed S, Zhang JK, Benedict B, Botterbush K, Yakdan S, Molina CA, Pennicooke BH, Hafez D, Ogunlade JI, Pallotta N, Gupta MC, Buchowski JM, Neuman B, Steinmetz M, Ghogawala Z, Kelly MP, Goodin BR, Piccirillo JF, Rodebaugh TL, Lu C, Ray WZ. Preoperative mobile health data improve predictions of recovery from lumbar spine surgery. Neurosurgery 2024;95:617.

[17] Hamaker EL, Grasman RPPP. Regime switching state-space models applied to psychological processes: Handling missing data and making inferences. Psychometrika 2012;77:400–422.

[18] Ji L, Chow S-M, Schermerhorn AC, Jacobson NC, Cummings EM. Handling missing data in the modeling of intensive longitudinal sata. Structural Equation Modeling: A Multidisciplinary Journal 2018;25:715–736.

[19] Jones A, Remmerswaal D, Verveer I, Robinson E, Franken IHA, Wen CKF, Field M. Compliance with ecological momentary assessment protocols in substance users: a meta-analysis. Addiction 2019;114:609–619.

[20] Knight SR, Ng N, Tsanas A, Mclean K, Pagliari C, Harrison EM. Mobile devices and wearable technology for measuring patient outcomes after surgery: a systematic review. NPJ Digital Medicine 2021;4:157.

[21] Leroux A, Crainiceanu C, Zeger S, Taub M, Ansari B, Wager TD, Bayman E, Coffey C, Langefeld C, McCarthy R, Tsodikov A, Brummet C, Clauw DJ, Edwards RR, Lindquist MA, ; A2CPS Consortium. Statistical modeling of acute and chronic pain patient-reported outcomes obtained from ecological momentary assessment. Pain 2024. doi:10.1097/j.pain.0000000000003214.

[22] Majumder S, Mondal T, Deen MJ. Wearable Sensors for Remote Health Monitoring. Sensors 2017;17:130.

[23] Mansueto AC, Wiers RW, van Weert JCM, Schouten BC, Epskamp S. Investigating the feasibility of idiographic network models. Psychological Methods 2022.

[24] May M, Junghaenel DU, Ono M, Stone AA, Schneider S. Ecological momentary assessment methodology in chronic pain research: A systematic review. The Journal of Pain 2018;19:699–716.

[25] McNeish D, Hamaker EL. A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. Psychological Methods 2019:No Pagination Specified-No Pagination Specified.

[26] Murray AL, Brown R, Zhu X, Speyer LG, Yang Y, Xiao Z, Ribeaud D, Eisner M. Prompt-level predictors of compliance in an ecological momentary assessment study of young adults' mental health. Journal of Affective Disorders 2023;322:125–131.

[27] Myin-Germeys I, Kasanova Z, Vaessen T, Vachon H, Kirtley O, Viechtbauer W, Reininghaus U. Experience sampling methodology in mental health research: new insights and technical developments. n.d.

[28] Myin-Germeys I, Oorschot M, Collip D, Lataster J, Delespaul P, Os J van. Experience sampling research in psychopathology: opening the black box of daily life. Psychological Medicine 2009;39:1533–1547.

[29] Ono M, Schneider S, Junghaenel DU, Stone AA. What Affects the Completion of Ecological Momentary Assessments in Chronic Pain Research? An Individual Patient Data Meta-Analysis. Journal of Medical Internet Research 2019;21:e11398.

[30] Osman A, Barrios FX, Gutierrez PM, Kopper BA, Merrifield T, Grittmann L. The Pain Catastrophizing Scale: Further Psychometric Evaluation with Adult Samples. Journal of Behavioral Medicine 2000;23.

[31] Perski O, Keller J, Kale D, Asare BY-A, Schneider V, Powell D, Naughton F, ten Hoor G, Verboon P, Kwasnicka D. Understanding health behaviours in context: A systematic review and meta-analysis of ecological momentary assessment studies of five key health behaviours. Health Psychology Review 2022;16:576–601.

[32] Pew Research Center. Demographics of Mobile Device Ownership and Adoption in the United States. Pew Research Center Internet & Technology 2019. Available: https://www.pewresearch.org/internet/fact-sheet/mobile/. Accessed 22 Apr 2020.

[33] Reiter T, Schoedel R. Never miss a beep: Using mobile sensing to investigate (non-)compliance in experience sampling studies. Behav Res 2024;56:4038–4060.

[34] Rogers AH, Rabbitts JA, Saper MG, Schmale GA, Palermo TM, Groenewald CB. Ecological momentary assessment of sleep, pain, and opioid use among adolescents following surgery. SLEEP Advances 2024;5:zpae039.

[35] Shiffman S, Stone AA, Hufford MR. Ecological Momentary Assessment. Annual Review of Clinical Psychology 2008;4:1–32.

[36] Sluka KA, Wager TD, Sutherland SP, Labosky PA, Balach T, Bayman EO, Berardi G, Brummett CM, Burns J, Buvanendran A, Caffo B, Calhoun VD, Clauw D, Chang A, Coffey CS, Dailey DL, Ecklund D, Fiehn O, Fisch KM, Frey Law LA, Harris RE, Harte SE, Howard TD, Jacobs J, Jacobs JM, Jepsen K, Johnston N, Langefeld CD, Laurent LC, Lenzi R, Lindquist MA, Lokshin A, Kahn A, McCarthy RJ, Olivier M, Porter L, Qian W-J, Sankar CA, Satterlee J, Swensen AC, Vance CGT, Waljee J, Wandner LD, Williams DA, Wixson RL, Zhou XJ, Consortium the A. Predicting chronic postsurgical pain: current evidence and a novel program to develop predictive biomarker signatures. PAIN 2023;164:1912.

[37] Solhan MB, Trull TJ, Jahng S, Wood PK. Clinical assessment of affective instability: Comparing EMA indices, questionnaire reports, and retrospective recall. Psychological Assessment 2009;21:425–436.

[38] St. Louis County, Missouri - Census Bureau Profile. n.d. Available: https://data.census.gov/profile/St._Louis_County,_Missouri?g=050XX00US29189#race-and-ethnicity. Accessed 24 July 2025.

[39] Stone AA, Broderick JE, Shiffman SS, Schwartz JE. Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain. Pain 2004;107:61–69.

[40] Stroop JR. Studies of interference in serial verbal reactions. Journal of Experimental Psychology 1935;18:643–662.

[41] Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: Development and Validation. Psychological Assessment 1995;7:524–532.

[42] Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. Journal of Clinical Epidemiology 2000;53:207–216.

[43] Tomczak M, Tomczak E. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. Biblioteka Akademii Wychowania Fizycznego w Poznaniu 2014. Available: https://www.wbc.poznan.pl/dlibra/publication/413565. Accessed 27 June 2025.

[44] Tonkin S, Gass J, Wray J, Maguin E, Mahoney M, Colder C, Tiffany S, Jr LWH. Evaluating Declines in Compliance With Ecological Momentary Assessment in Longitudinal Health Behavior Research: Analyses From a Clinical Trial. Journal of Medical Internet Research 2023;25:e43826.

[45] Trull TJ, Ebner-Priemer UW. Ambulatory assessment in psychopathology research: A review of recommended reporting guidelines and current practices. Journal of Abnormal Psychology 2020;129:56–63.

[46] Turner CM, Coffin P, Santos D, Huffaker S, Matheson T, Euren J, DeMartini A, Rowe C, Batki S, Santos G-M. Race/ethnicity, education, and age are associated with engagement in ecological momentary assessment text messaging among substance-using MSM in San Francisco. Journal of Substance Abuse Treatment 2017;75:43–48.

[47] Vachon H, Viechtbauer W, Rintala A, Myin-Germeys I. Compliance and retention with the experience sampling method over the continuum of severe mental disorders: Meta-analysis and recommendations. Journal of Medical Internet Research 2019;21. doi:10.2196/14475.

[48] Votaw VR, Witkiewitz K. Motives for Substance Use in Daily Life: A Systematic Review of Studies Using Ecological Momentary Assessment. Clinical Psychological Science 2021;9:535–562.

[49] Wells JE, Horwood LJ. How accurate is recall of key symptoms of depression? A comparison of recall and longitudinal reports. Psychological medicine 2004;34:1001–1011.

[50] Williams MT, Lewthwaite H, Fraysse F, Gajewska A, Ignatavicius J, Ferrar K. Compliance With Mobile Ecological Momentary Assessment of Self-Reported Health-Related Behaviors and Psychological Constructs in Adults: Systematic Review and Meta-analysis. Journal of Medical Internet Research 2021;23:e17023.

[51] Wrzus C, Neubauer AB. Ecological Momentary Assessment: A Meta-Analysis on Designs, Samples, and Compliance Across Research Fields. Assessment 2023;30:825–846.

[52] Yao L, Yang Y, Wang Z, Pan X, Xu L. Compliance with ecological momentary assessment programmes in the elderly: a systematic review and meta-analysis. 2023. doi:10.1136/bmjopen-2022-069523.