



# Augmenting the Clinical Data Sources for Enigmatic Diseases: A Cross-Sectional Study of Self-Tracking Data and Clinical Documentation in Endometriosis

Ipek Ensari<sup>1</sup> Adrienne Pichon<sup>2</sup> Sharon Lipsky-Gorman<sup>2</sup> Suzanne Bakken<sup>2,3</sup> Noémie Elhadad<sup>2</sup>

<sup>1</sup>Data Science Institute, Columbia University, New York, New York, United States

<sup>2</sup>Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, New York, United States

<sup>3</sup>Columbia School of Nursing, Columbia University, New York, New York, United States

**Address for correspondence** Ipek Ensari, PhD, Columbia University, Data Science Institute, 475 Riverside Dr, Room 320, New York, NY 10115, United States (e-mail: ie2145@columbia.edu).

Appl Clin Inform 2020;11:769–784.

## Abstract

**Background** Self-tracking through mobile health technology can augment the electronic health record (EHR) as an additional data source by providing direct patient input. This can be particularly useful in the context of enigmatic diseases and further promote patient engagement.

**Objectives** This study aimed to investigate the additional information that can be gained through direct patient input on poorly understood diseases, beyond what is already documented in the EHR.

**Methods** This was an observational study including two samples with a clinically confirmed endometriosis diagnosis. We analyzed data from 6,925 women with endometriosis using a research app for tracking endometriosis to assess prevalence of self-reported pain problems, between- and within-person variability in pain over time, endometriosis-affected tasks of daily function, and self-management strategies. We analyzed data from 4,389 patients identified through a large metropolitan hospital EHR to compare pain problems with the self-tracking app and to identify unique data elements that can be contributed via patient self-tracking.

**Results** Pelvic pain was the most prevalent problem in the self-tracking sample (57.3%), followed by gastrointestinal-related (55.9%) and lower back (49.2%) pain. Unique problems that were captured by self-tracking included pain in ovaries (43.7%) and uterus (37.2%). Pain experience was highly variable both across and within participants over time. Within-person variation accounted for 58% of the total variance in pain scores, and was large in magnitude, based on the ratio of within- to between-person variability (0.92) and the intraclass correlation (0.42). Work was the most affected daily function task (49%), and there was significant within- and between-person variability in self-management effectiveness. Prevalence rates in the EHR were significantly lower, with abdominal pain being the most prevalent (36.5%).

**Conclusion** For enigmatic diseases, patient self-tracking as an additional data source complementary to EHR can enable learning from the patient to more accurately and comprehensively evaluate patient health history and status.

## Keywords

- mobile platforms
- quantified self
- chronic illness or special needs
- electronic health records and systems
- personal health records
- smart phone

received  
May 17, 2020  
accepted  
July 14, 2020

**DOI** <https://doi.org/10.1055/s-0040-1718755>.  
**ISSN** 1869-0327.

© 2020 Georg Thieme Verlag KG  
Stuttgart · New York

**License terms**



## Background and Significance

The Learning Health System (LHS) framework proposes an improved health care system to accelerate the lengthy (~17 years) learning health cycle from health care data to research findings that inform clinical decisions by learning and sharing information from patients, care delivery systems, and clinical research over time.<sup>1,2</sup> This framework relies on the electronic health record (EHR) as a primary source for gaining information about patients and generating evidence to inform care,<sup>3</sup> and aims to make the research-to-clinical care pipeline more efficient<sup>4</sup> through use of big data, health technologies, and participatory design.<sup>5,6</sup>

The EHR constitutes a rich, convenient source of data for clinical decision-making,<sup>7</sup> conducting biomedical research,<sup>8,9</sup> and generating relevant and actionable knowledge that can inform the tailoring of treatments.<sup>10</sup> The problem lists within the EHR are particularly useful as they provide a structured, comprehensive, and accessible list of patient problems.<sup>11</sup> Problem lists include diagnoses, illnesses, injuries, and other factors that affect the health of the patient, documented by the provider at the point of care and stored in standard encoding systems.<sup>12</sup> They facilitate data sharing, information retrieval, and communication throughout the health care continuum, and further provide data source for research studies and other secondary data uses.<sup>12</sup>

Yet, sole reliance on the EHR and data collected at point of care can be insufficient,<sup>13</sup> particularly when studying conditions that lack adequate documentation,<sup>14</sup> conditions that have a fluctuating course over time, and in instances where the clinic assessment of an outcome might differ from its ambulatory assessment.<sup>15</sup> Health providers may not log diagnosis codes for observed conditions that do not require treatment, diagnostic testing, or referral.<sup>8</sup> Previous studies further report incompleteness and significant variability in provider agreement on problem lists,<sup>16</sup> and significant burnout due to EHR use among providers<sup>17–19</sup> resulting in less time treating patients, incomplete chart notes, and duplicate records.<sup>17</sup> These factors can collectively contribute to deficiencies in EHR data accuracy, quality, and level of detail.<sup>20</sup> This can subsequently pose challenges for not only clinical care,<sup>21</sup> but also research including disease case identification<sup>14,22</sup> and longitudinal analyses.<sup>20</sup> In the context of enigmatic diseases, these deficiencies can further challenge attempts to understand and describe (i.e., “characterize”) the disease, derive accurate and comprehensive evaluations of patient’s disease status and subsequently target their needs.

One approach to circumvent these limitations is to augment the EHR data with direct patient input obtained through self-tracking.<sup>23–25</sup> Self-tracking is defined as the practice of repeatedly collecting data and reflecting on it to acquire knowledge or achieve a goal such as behavior change.<sup>26–29</sup> Since these data capture patients’ day-to-day experiences outside of formal clinical settings,<sup>28,30</sup> they can aid in differential diagnosis<sup>31</sup> and monitoring of disease by providing additional information not obtained from the EHR. In addition to contributing to better patient care and shared clinical decision-making,<sup>24,32–34</sup> including the patient in the

data generation and sharing process further circumvents placing additional documentation burden on the provider at point of care. As such, this approach could be an attractive alternative to efforts to increase provider EHR documentation responsibility.<sup>35</sup> The potential impact of self-tracked health data through mHealth technology for informing evidence-based care is acknowledged by many,<sup>36,37</sup> including the World Health Organization (WHO) stating that it can help address the growing burden of chronic disease.<sup>38</sup> These factors collectively provide impetus for investigating the data elements that can be elicited from mobile patient tracking data for describing enigmatic diseases.

One such enigmatic condition is endometriosis: a systemic, estrogen-dependent inflammatory condition with a range of debilitating symptoms,<sup>39</sup> comorbidities,<sup>40</sup> and associated with significant impact on daily function and quality of life (QoL).<sup>41,42</sup> Chronic pelvic pain is its most frequent symptom<sup>43</sup>; others include pain with sexual intercourse (dyspareunia), painful urination (dysuria), and ovulation pain.<sup>44,45</sup> Endometriosis is associated with a productivity loss of 6.3 hours/week<sup>46</sup> and an estimated \$69.4 billion per year in excess health expenditures in the United States,<sup>42</sup> and is the second leading indication for hysterectomy.<sup>47</sup> Despite its prevalence rate of 10% among women of reproductive age,<sup>44</sup> endometriosis is an underfunded disease<sup>48</sup> that is still considered “clinically confusing” and “surprising,”<sup>44,49</sup> with no cure or adequate treatment to date. There further is substantial between-patient variation in its clinical manifestations,<sup>44,49</sup> which likely contributes to the approximately 6.7 (± 6.3) year delay between symptom-onset and its surgical diagnosis,<sup>50</sup> and the difficulty of understanding its heterogeneous pathophysiology and identification of possible subtypes (i.e., “phenotypes”) of the disease.

As a disease demographic, endometriosis patients further face inequities with respect to access to clinical care,<sup>40,51</sup> including inadequate insurance coverage of endometriosis-related treatments,<sup>40,52</sup> and gender bias and attitudes influencing physician’s diagnostic and treatment decisions,<sup>53,54</sup> which could be contributing to the underdocumentation of the patient’s disease experience. An investigation<sup>55</sup> of the patterns of prevalence, incidence, and frequency of endometriosis reported that no data are available across the past 30 years to attempt to document changes in presenting symptom severity, life impact, or short- or long-term prognosis among women with endometriosis. Failure to obtain the right data about these patients is a missed opportunity for effectively targeting their unique needs,<sup>56</sup> and we propose that this lack of proper documentation can be augmented through patient self-report. Improving documentation can provide the foundation to address the inequities in diagnosis and care for those with endometriosis and that including the patient in the generation of such documentation can be a starting point in this endeavor.

Our inquiry is in line with the goal of LHS to create patient-centered care systems and is further supported by previous recommendations<sup>57</sup> for leveraging health information technologies to support and implement this goal. It is proposed that patient-centered information technologies can be leveraged to democratize the health care system within the LHS

through patient engagement and inclusion in knowledge co-creation. As such, our study addresses a timely question by exploring the potential of patient self-tracking for augmenting the currently relied upon clinical data sources.

## Objectives

Using endometriosis as a grounding example, we investigate the additional information that can be gained through mHealth self-tracking on conditions that are poorly understood and documented. Accordingly, this study aims to (1) describe endometriosis pain problems documented through self-tracking and in the EHR to investigate the potential of self-tracking for augmenting the EHR documentation, (2) quantify the between- and within-person fluctuations in pain to characterize the dynamic nature of endometriosis, and (3) assess endometriosis impact on daily function and its self-management documented through self-tracking as additional data elements that can provide information about the burden of disease and possible points of intervention for disease management.<sup>42,43</sup>

## Methods

Study design and protocols were approved by the Columbia University Irving Medical Center (CUIMC) Institutional Review Board (AAAQ9812). This was a retrospective study conducted with data from two independent samples of women with a clinician or surgery confirmed diagnosis of endometriosis.

### Self-Tracking Sample

We obtained self-tracking data from Phendo, an observational research app for tracking endometriosis symptoms and available for iOS<sup>1</sup> and Android.<sup>2</sup> Phendo allows anyone with endometriosis around the world to download the app at any time to characterize their endometriosis experience by tracking symptoms, activities, and self-management techniques. Users 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. Endometriosis diagnosis is determined through self-report based on participants' response to a question asking whether the participant has endometriosis and if so, how it was diagnosed (i.e., surgery or clinician confirmation). Phendo was designed using a participatory design approach through a series of qualitative and quantitative studies with endometriosis patients, with the goal of creating a patient-centered tool that engages the user as an active participant in the research on better understanding endometriosis.<sup>58–60</sup> 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. Aligned with goals of the LHS for promoting patient engagement,<sup>6,61</sup> these aspects of Phendo make it particularly suitable for undertaking this investigation.

### Electronic Health Record Cohort Extraction

The Observational Health Data Sciences and Informatics (OHDSI) de-identified instance of the CUIMC Data Warehouse<sup>62,63</sup> was queried to select a cohort of endometriosis

cases for description of pain problems documented in the EHR. The CUIMC EHR system is used across the medical center university hospitals, clinics, and doctors' offices across the City of New York, which has a population of approximately 8.44 million, including 43% White, 24% Black, 14% Asian, 15% other race, and 3.5% two or more races.<sup>64</sup> The clinical data warehouse includes both inpatient and outpatient records and contains 36,578 "concepts" (i.e., clinical entities such as findings, diagnoses, drugs, and procedures) including 11,952 conditions, 12,334 drugs, and 10,816 procedures from 5,364,781 patients.<sup>65</sup>

We extracted the data using the OHDSI Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM)<sup>62,66,67</sup> and SNOMED CT as the standardized EHR vocabulary.<sup>68</sup> The CDM formally specifies the encoding and relationships among EHR concepts in a consistent and standardized way using a hierarchy of subtypes.<sup>62,68</sup> This harmonizes disparate coding systems with minimal information loss to a standardized vocabulary by mapping source concepts from different systems (e.g., ICD-9, ICD-10, RxNorm, CPT4, NDC, etc.) onto a standard concept ID during the extract-transform-load (ETL) process. SNOMED CT is the designated U.S. standard terminology for EHR diagnosis and problem lists, and has distinct advantages over ICD codes for capturing discrete patient care diagnoses.<sup>69,70</sup> It hierarchically organizes concepts, which allows aggregation of information based on subtype classification. More specific (i.e., descendent or child) concepts have more granularity and detail, whereas more general (i.e., ascendent or parent) concepts have less granularity and clinical detail, but they aggregate similar subtype concepts (e.g., "abdominal tenderness of left lower quadrant" and "right lower quadrant pain") into logical groups (e.g., "abdominal pain") which can be a desirable feature when conducting statistical analyses.

### Electronic Health Record Sample

A cohort of women with endometriosis in the EHR was created using the following set of criteria that have previously been demonstrated<sup>71</sup> to have a specificity of 93% and precision of 85% for correctly selecting endometriosis patients in the EHR: women 15 to 49 years old with an endometriosis-related or endometriosis-prevalent procedure, an endometriosis concept code 30 days prior or post either of these procedures and documentation of an endometriosis concept code after the initial procedure. For endometriosis condition codes, we used the SNOMED concepts "endometriosis" (129103003), "endometriosis of the ovary" (266589005), "endometriosis of the pelvic peritoneum" (198251001), and "endometriosis of the uterus" (76376003).

## Outcomes

### Self-Track Pain Problems

The Phendo app includes four pain-related questions (→ **Table 1**). Item "Where is the pain?" allows users to select responses (e.g., "cervix," "vagina," and "chest") using a visual pain scale (→ **Supplementary Figure S1**, available in the online version), similar to the McGill Pain Questionnaire.<sup>72</sup>

**Table 1** Phendo pain-related variables and tracking statistics ( $n = 6,925$ )

Question	Sample size	Tracked instances (median/mean/SD/range)	Tracking frequency in days (mean/SD/range)
Where is the pain? How severe is the pain? <sup>a</sup>	5,205	42/105.5/162.5/1–965	6.3/26.4/1–1,147
Any other pain symptoms? <sup>b</sup> How severe is the symptom? <sup>a</sup>	4,413	25/66.6/103.8/1–536	8.2/30.7/1–1,147
Any gastrointestinal/urine issues? How severe is the symptom? <sup>b</sup>	4,578	30/71.9/106.8/1–573	7.4/29.1/1–1,147
How was sex? (dyspareunia)	676	50/85.0/87.5/1–344	4.7/19.7/1–609
Which activities are hard to do?	5,764	36/96.5/142.1/1–708	5.5/28.2/1–999
What did you do to self-manage? Did it help? (“helpful” or “not helpful/no effect”)	6,025	40/106.0/164.9/1–958	5.3/27.4/1–1,019

Abbreviation: SD, standard deviation.

<sup>a</sup>Severity of the pain is tracked as “mild” (1), “moderate” (2), and “severe” (3).

<sup>b</sup>Responses to item “any other pain symptoms?” were included to capture pain reports in instances where a participant might have tracked this item but not item 1 or 3.

Item “any gastrointestinal/urine symptoms?” allows participants to report painful bowel movements (i.e., dyschezia), dysuria, and epigastric (e.g., gas) pain.

### Pain Problems in the Electronic Health Record

We selected all pain-related problems documented in the patients’ records and computed their prevalence rates to describe documentation of endometriosis-related pain in the EHR. We chose to investigate the problem lists in the EHR due to their relevance for both research and clinical practice. Problem lists provide a structured, systematic way to contain clinician notes, facilitating access to data and conducting research.<sup>11</sup> From a clinical standpoint, they improve efficiency for multiple providers to coordinate patient management, and for those who are taking care of new patients to familiarize themselves with a patient’s problems quickly.<sup>73</sup>

### Pain Variability

Investigations of within- and between-individual variability in endometriosis pain were conducted using self-tracked data from Phendo participants. While significant pain variability has previously been reported in other conditions,<sup>74</sup> between- versus within-individual variability in pain over time in endometriosis has not been quantified. We undertook this investigation because disease course over time and across patients can have important implications for deriving comprehensive profiles of patient disease history and status, identification of disease phenotypes,<sup>74</sup> and for predicting change in disease course or functional capacity over time.<sup>75</sup>

To jointly assess variability for both dimensions of area and severity of pain, we created a composite day-level pain score 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), removing any problem-severity duplicate for the day. Removing day-level duplicates for pain area-severity pairs circumvents several issues: (1) the possibility of rumination/catastrophizing, which might look like, for example, a participant tracking

the same intensity of pain in the same area multiple times within a span of several hours, (2) to account for the day-level tracking habits of the participants (e.g., some participants might track a pain area once at the end of the day even if they experience it multiple times during the day), and (3) any possible errors during data transfer from the app to the data warehouse (e.g., if there is no reception or wi-fi at the time of tracking, there might be a delay in data dumping, potentially though infrequently leading to duplicates of the entry into the data warehouse).

### Additional Data Elements

Prevalence of affected daily function tasks and self-management techniques were assessed for the Phendo sample. Participants can track free-text responses to these two items, which then get mapped onto common terms for standardization (e.g., “go to a family gathering,” “hang out with friends” get mapped onto “socialize”). This standardization process has been used in other previous research from our group,<sup>71,76</sup> which we maintain for this analysis, and rely on published literature<sup>43,52</sup> for generalizability and comparability of the prevalence estimates.

### Data Analytic Approach

#### Prevalence of Pain Problems

In Phendo, we computed total counts and percentages of pain problems across the entire sample by including the total sample size in the denominator instead of limiting to the number of participants to those who ever tracked a pain problem, as a more conservative approach. In the EHR sample, we computed counts and percentages of pain problems, aggregating at the ascendent concept level, as this allows for logical groupings of pain problems.

For pain problems that exist in both samples, we conducted two-sample tests for equality of proportions.<sup>77</sup> We further computed the standardized  $z$  scores and the confidence intervals around the difference between the two



proportions.<sup>78</sup> We a priori chose to compute prevalence rates based on the aggregate parent concepts due to reasons related to the EHR data architecture described earlier (See *Electronic Health Record Cohort Extraction*).

### Variability of Pain

To describe the distribution of severity levels across pain problems at the sample level, we computed frequencies of three severity levels for each body area (i.e., each pain location-severity combination is summed and divided by the participant's total counts of tracks to account for tracking frequency). To quantify pain variability, we first computed the ratio of within-person to between-person variability<sup>75</sup> (i.e., average within-person standard deviation (SD) divided by the SD of the overall mean of all the participant-level mean scores). This ratio provides a standardized expression of the temporal variability relative to the distribution of the sample scores.<sup>75</sup> This is analogous to computing an effect size (e.g., Cohen's *d* or a *t*-score) where the magnitude of effect is expressed in SDs, and provides two advantages that make it particularly suitable when there is both substantial average within- and between-person variability in the data. First, because change is expressed as a standardized score, change and its statistical significance can be evaluated for an individual without referring to scores from other individuals. Second, it incorporates between-person differences in magnitude of variability by adjusting the change in terms of each individual's magnitude of variability.<sup>75</sup>

Next, we estimated the between- and within-person variances using a linear mixed-effects model (LMM) where day-level pain was regressed on the grouping variable of participant as a random effect. LMMs are commonly used in repeated measures designs and provide estimates of variance in the outcome (i.e., pain) explained by the predictor variables (i.e., participant).<sup>79</sup> The predictor term for participant in the model is included as a random effect, which handles nonindependence in the data and estimates a separate intercept for each participant. This provides an estimate of the between-participant variance, that is, variation in pain explained by the "grouping structure" where, group here refers to each participant. In this type of model where the only predictor is the participant variable and the intercept, the total variance is assumed to equal to the sum of between-participant and within-participant (i.e., the residual) variances. Using these estimates, we computed the intraclass correlation (ICC), which is calculated as the variance among participant means over the sum of participant-level and data-level (residual) variances.<sup>80</sup> The ICC (also referred to as "repeatability") is a common and recommended measure for assessing stability or fluctuation over time, and represents the fraction of the total variance in the population of interest that can be attributed to variation among groups.<sup>80</sup> It further provides an estimation of the expected correlation between two randomly drawn units that are in the same group<sup>81</sup> (e.g., two repeated measurements of pain for one individual). Larger ICC values indicate larger within-group correlation or repeatability and larger magnitude of the variance explained by the between-group variation. Finally, we estimated the confidence intervals for the ICC

using parametric bootstrapping and the statistical significance of the ICC in the LMM using a likelihood ratio test (LRT).<sup>80,82</sup> The LRT assesses goodness-of-fit between two competing (e.g., null and alternative) models by comparing their log likelihoods.

### Additional Data Elements

Using the Phendo sample, we computed prevalence rates for daily function tasks, self-management techniques, and frequency of effectiveness (and its SD) of each self-management technique.

## Results

### Sample Characteristics

Phendo sample characteristics are provided in **Table 2**. There were 6,925 participants with a surgical- or clinician-confirmed diagnosis of endometriosis who responded to at least one of the questions pertaining to pain, daily function tasks, or self-management techniques. The EHR query identified 4,389 endometriosis patients, which translates to a prevalence of 0.4% documented endometriosis cases (i.e., out of 1,097,250 females in the same age bracket documented within the same time frame) in the clinical database. The sample consisted of women aged 14 to 49, with a mean age of 36.4 (SD = 7.4) and median age of 37 (mean absolute deviation = 7.4). With respect to race, 1,227 (56.0%) were white, 431 (9.8%) were Black/African-American, 202 (4.6%) were Asian, 18 (0.4%) were other Pacific-Islander, 23 (0.5%) were Asian Indian, 3 (0.06%) were American Indian or Alaska Native, 20 (0.4%) were of other race, and 2,462 (56.0%) did not have race data. Ethnicity data were unavailable on more than half the patients; 651 (14.8%) were Hispanic, 1,420 (32.3%) were non-Hispanic, and no data were available from 2,318 patients. We further searched for data on other demographic factors (e.g., such as employment and education status); however, no further information was available.

### Outcomes

#### Description of Pain Problems in Phendo

A visual depiction of the prevalence of pain problems in Phendo are provided in **Figure 1**. Pelvic pain was the most prevalent problem (57.3%), followed by lower back pain (49.2%) and ovarian pain (43.7%). Gastrointestinal (i.e., epigastric and intestinal) problems collectively were prevalent in 55.9% of the sample, comparable to that of pelvic pain. Unique pain problems identified through self-tracking included pain in ovaries, uterus, vagina, cervix, dyschezia, dyspareunia, and rectal pain (**Supplementary Table S1**, available in the online version).

#### Description of Pain Problems in the Electronic Health Record

Pain problem counts and prevalence rates in the EHR sample are provided in **Supplementary Table S2** (available in the online version). We identified 19 parent pain concepts and 101 unique descendent concepts mapped onto one or more parent concepts. Abdominal pain was the most prevalent

**Table 2** Descriptive statistics for the Phendo sample

Participant characteristics	Mean (SD)/Frequency (%)
Age (y) (N = 5,915)	29.6 (6.9) Median = 28.9 (MAD = 7.2)
Body mass index (N = 5,522) <sup>a</sup>	26.5 (7.0) Median = 24.7 (MAD = 5.6)
Endometriosis diagnosis	
Surgically confirmed	5,224 (75.4)
Clinician confirmed	1,701 (24.5)
Employment status	
Employed	3,876 (55.9)
Not employed	856 (12.3)
Student	844 (12.1)
Unknown	1,349 (19.4)
Race/Ethnicity	
Non-Hispanic White	4,861 (70.1)
Non-Hispanic Black	160 (2.3)
Hispanic (White or non-White)	327 (4.7)
Asian	156 (2.2)
Native American	45 (0.6)
Other (including mixed race/ethnicity)	371 (5.3)
Unknown	1,005 (14.5)
Marital status	
Married/in a domestic partnership	3,298 (47.5)
Separated/divorced	191 (2.7)
Single, never married	2,086 (30.1)
Widowed	3 (0.04)
Unknown	1,016 (14.6)
Living environment	
Urban	2,428 (35.0)
Suburban	2,489 (35.9)
Rural	992 (14.3)
Unknown	17 (0.38)
Primary work location	
Home	1,481 (21.3)
Outside home	4,084 (58.9)
Unknown	1,360 (19.6)
Education level	
High school or less	831 (2)
Some college	1,828 (26.3)
College or higher	3,251 (46.9)
Unknown	1,015 (14.6)

Abbreviations: BMI, body mass index; MAD, mean absolute deviation; SD, standard deviation.

Note: Unknown indicates not tracked by the participant in the Phendo app profile, and therefore, data are not available.

<sup>a</sup>Values below 12 and over 70 are excluded from BMI summary statistics.

pain problem (36.5%), followed by pain in pelvis (29.8%). Rest of the pain problems were documented for less than 10% of the patients, including chest pain (9%), dysuria (4.3%), and backache (6.8%). Of note, 856 pelvic pain concepts were also mapped under “abdominal pain,” which included dysmenorrhea, renal colic, and bladder pain. Excluding these repeated cases, there were 1,147 unique “abdominal pain” concepts documented in the EHR sample (26.1%). Upon further inspection, we identified 101 instances of breast pain under the “chest pain” concept (2.3%), and 128 instances of epigastric pain under “abdominal pain” (2.9%).

### Two Samples Proportion Tests

Results from the two-sample proportion tests are provided in **Table 3**. All pain problems were significantly more prevalent in the Phendo sample ( $p = 0.01$  for abdominal pain,  $p < 0.001$  for all others). Largest differences in proportions were observed for back pain, headache/migraines, and pelvic pain (**Table 3**).

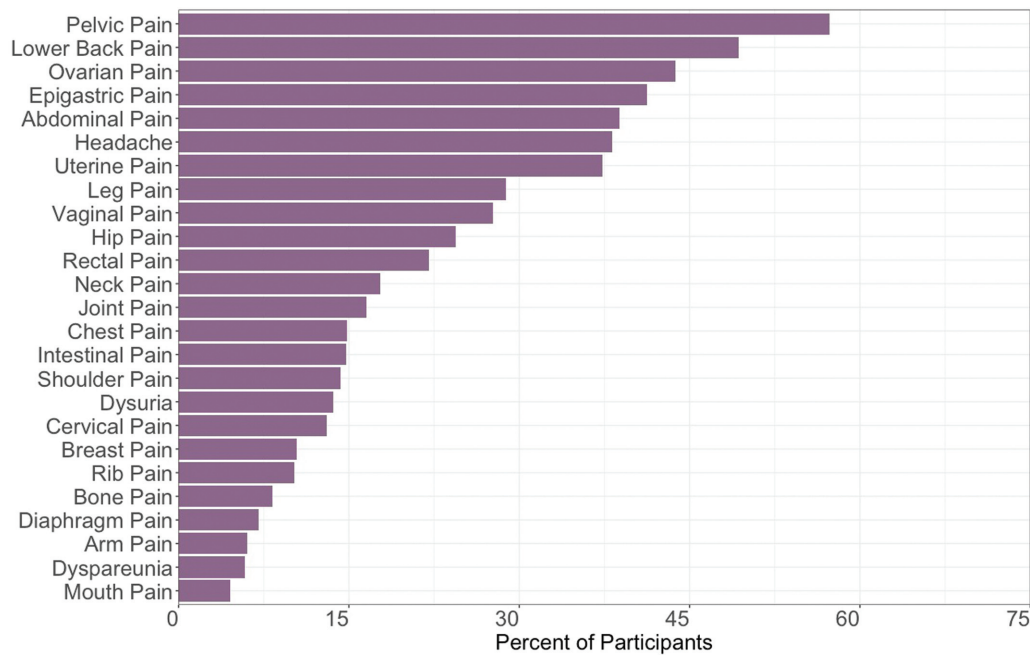
### Variability in Pain

#### Qualitative Description of Variability

Visualization of self-tracked pain reports at different levels of severity and through time indicated large variability in the average intensities reported across participants (**Figures 2 and 3**). Moderate pain was the most prevalent severity across body areas (**Figure 2** top), in line with previous literature on pain typology in endometriosis.<sup>45</sup> However, tracking frequencies of all three severity levels significantly varied for all pain problems based on the SDs (**Figure 2** bottom). The trajectories of pain frequencies in a subsample of 194 participants over a 4-week duration are depicted using a Sankey's diagram (**Figure 3** top), which is a type of flow diagram for showing factors' states and transitions over time.<sup>83</sup> For easier visual representation of the differential path transitions represented by the “links” across 4 weeks, weekly frequencies are categorized as low (<5 times), moderate (5–10 times), or high (>10 times). The cut-off points were selected such that each 5-point interval corresponds to approximately half a SD, based on the sub-sample's mean (9.0) and SD (11.2) for weekly pain frequency. The links connecting the bars (“nodes”) depict the variability in the pain frequency through time, both within and between participants. The tile grid diagram (**Figure 3** bottom) shows daily fluctuations in pain location and severity over time for a single Phendo participant. Darker colors indicate larger number of pain locations reported by the participant. For example, the depicted participant reported severe pain in three body locations on March 19, while on March 20, they reported moderate and severe pain in six body locations each.

#### Quantitative Description of Variability

All participants who provided at least 7 days of tracking data on pain were included in the variability analysis, yielding a total of 43,539 person-level days from 1,534 Phendo participants. Mean average (i.e., “mean of means”) day-level pain score was 7.75 (SD = 5.61). Median average day-level pain score was 6.21 (mean absolute deviation = 3.79). Within-



**Fig. 1** Prevalence of pain problems reported in the Phendo sample.

**Table 3** Results of the two-sample test for equality of proportions

Pain problem (SNOMED ID)	EHR counts (proportion)	Phendo counts (proportion)	Chi-square statistic	95% CI	z statistic
Abdominal <sup>a</sup> (200219)	1,603 (0.36)	2,690 (0.38)	6.14 ( $p = 0.01$ )	−0.04 to 0.004	−2.47
Abdominal (200219)	885 (0.20)	2,690 (0.38)	433.74 <sup>b</sup>	−0.20 to 0.17	−20.82
Headache/Migraines (378253/318736)	480 (0.05)	2,644 (0.38)	1,494 <sup>b</sup>	−0.33 to 0.31	−38.65
Backache (134736)	299 (0.06)	3,414 (0.49)	2,198.30 <sup>b</sup>	−0.43 to 0.41	−46.88
Chest (77670)	405 (0.09)	1,026 (0.14)	75.93 <sup>b</sup>	−0.06 to 0.04	−8.71
Joint (77074)	334 (0.07)	733 (0.10)	27.83 <sup>b</sup>	−0.04 to 0.02	−5.27
Dysuria (197684)	191 (0.04)	944 (0.13)	256.33 <sup>b</sup>	−0.10 to 0.08	−16.01
Lower limb (4024561)	261 (0.05)	1,998 (0.28)	882 <sup>b</sup>	−0.24 to 0.21	−29.69
Upper limb (4009890)	149 (0.03)	419 (0.06)	39.73 <sup>b</sup>	−0.03 to 0.06	−6.30
Hip (200219)	35 (0.003)	570 (0.08)	338.49 <sup>b</sup>	−0.08 to 0.07	−18.39
Bone (4129418)	16 (0.03)	570 (0.08)	338.49 <sup>b</sup>	−0.003 to 0.08	−18.39
Pelvis (4147829) <sup>a</sup>	1,312 (0.29)	3,969 (0.57)	811.57 <sup>b</sup>	−0.29 to 0.25	−28.48

Abbreviations: CI, confidence interval; EHR, electronic health record.

Note: 95% confidence interval built around the difference in proportions.

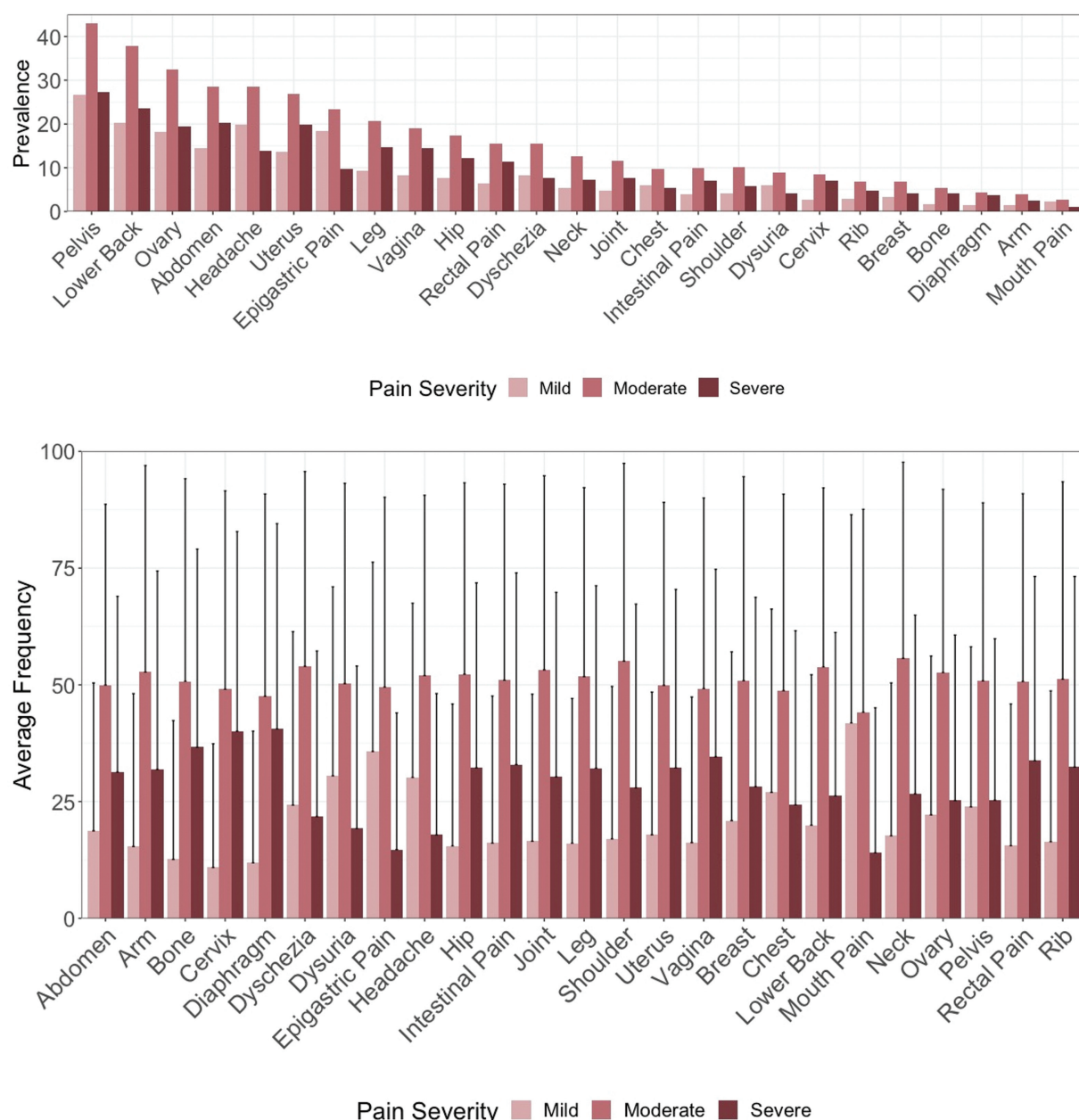
<sup>a</sup>Includes the 885 conditions that are mapped under both “abdominal pain” and “pain in pelvis.”

<sup>b</sup> $p < 0.0001$ .

person SD in day-level pain was 5.17 ( $SD = 3.45$ ). Expressed in SDs to obtain a standardized measure of the within-person variability<sup>75</sup> ( $5.17/5.61$ ), this yields a ratio of 0.92 and similar to other effects sizes, can be interpreted as large.<sup>75</sup>

Results of the LMM analysis are provided in **Table 4**. The ICC was 0.42 (95% confidence interval = 0.40–0.44) and was statistically significant (LRT Chi-square = 22,600,  $p < 0.001$ ). This indicates that the 42% of the variance in daily pain is attributed to between-participant variation and that pain scores are significantly correlated from day to day for each

participant. Within-person variance estimate was 38.72, indicating that approximately 58% of the total variance in daily pain is attributed to within-participant variation (**Table 4**). A plot of each participant's estimated mean scores and confidence intervals are provided in **Figure 4**, demonstrating the variability that occurs across participants (x-axis), and within-participant (y-axis). Taken together, these suggest that substantial variability occurs not just between participants but also within-participant over time. Of note, number of tracked days was not a significant



**Fig. 2** Top: Prevalence of mild, moderate, and severe pains by body location as reported in the Phendo cohort. Each body area-intensity pair counted once per participant. For all body locations, participants tracked predominantly moderate pain, with severe pain as second most common, and mild as least common. Bottom: Mean proportion of severity for each body area normalized by total tracks of pain reports and averaged across the sample. One-sided error bars were used for visual display purposes and represent standard deviations.

covariate in the variance estimation model ( $B = 0.001$ , standard error = 0.003,  $t = 0.485$ ,  $p = 0.62$ ), or correlated with pain scores (Pearson's  $r = 0.01$ ,  $t = 0.41$ ,  $p = 0.68$ ). Therefore, it was not included in the final model as a covariate.

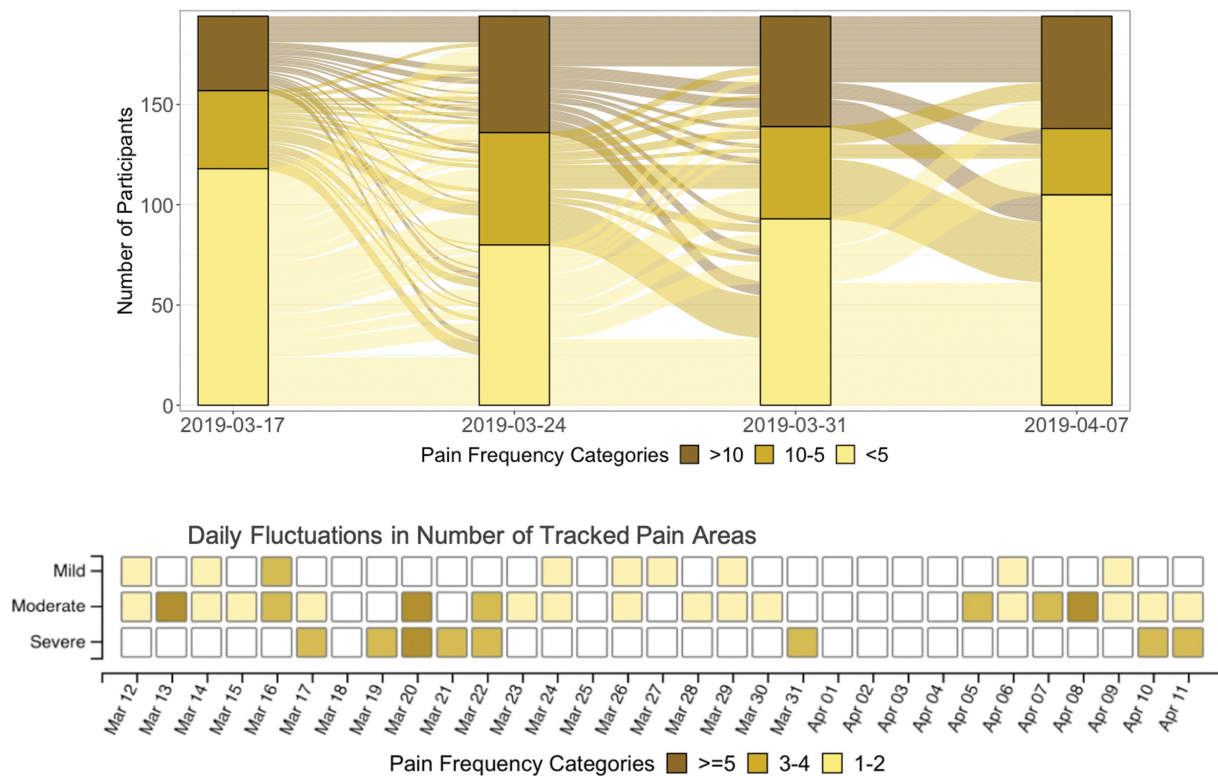
#### Additional Data Elements

Prevalence rates of reported affected daily function tasks are provided in [Table 5](#). We identified a total of 22 distinct tasks reported by the participants, who on average reported three unique affected tasks ( $SD = 2.2$ , range = 1–14). Working (53.1%), getting out of bed (52.2%), standing (47.4%), and using the toilet (47.3%) were the most frequently reported tasks. Of note, 43% of the participants reported sleeping as an affected

task. Similarly, there were multiple physical mobility-related tasks reported including sitting, walking, and climbing the stairs, reported by at least 30% of the participants.

Prevalence rates of all self-management techniques are provided in [Figure 5](#) and [Supplementary Table S3](#) (available in the online version). Participants reported a median of three unique techniques (mean = 3.21,  $SD = 2.21$ , range = 1–14). Rest and using a heat pack were the most prevalent (52.4 and 50.7%, respectively). Assessment of effectiveness frequency indicated considerable variability in the reported helpfulness of the techniques. For example, breathing exercises (reported as the third most frequently used technique by 38%) and ice packs were tracked as unhelpful 51 and 52% of





**Fig. 3** Top: Sankey diagram of pain frequencies over time in a sample of 194 Phendo participants over the course of a 4-week period. Weekly frequencies are categorized into low (<5 times), moderate (5–10 times), and high (>10 times), depicted by the different colors. The links depict the variability in the pain frequency through time, both within participant, and between participants. Bottom: Daily fluctuations in pain intensity over time for a sample individual in the Phendo cohort by severity and number of pain locations. Darker color indicates larger number of pain locations reported by the participant. For example, the depicted participant reported severe pain in three body locations on March 19, while on March 20, they reported moderate and severe pain in six body locations each.

**Table 4** Between- and within-person variance estimates obtained from the linear mixed-effects model ( $n = 1,534$ )

	Between-person variance ( $t^2$ )		Within-person variance ( $\sigma^2$ )	
	Variance	SD (95% CI)	Variance	SD (95% CI)
Pain score	28.7	5.3 (5.1–5.5)	39.2	6.2 (6.1–6.2)

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient; SD, standard deviation; SE, standard error.

Note: Phendo participants who provided at least 7 days of data were included in the analysis. Number of days was not a significant predictor in the model ( $B = 0.001$ ,  $SE = 0.003$ ,  $t = 0.485$ ,  $p = 0.62$ ) and therefore excluded from the final models for estimating variances and the ICC.

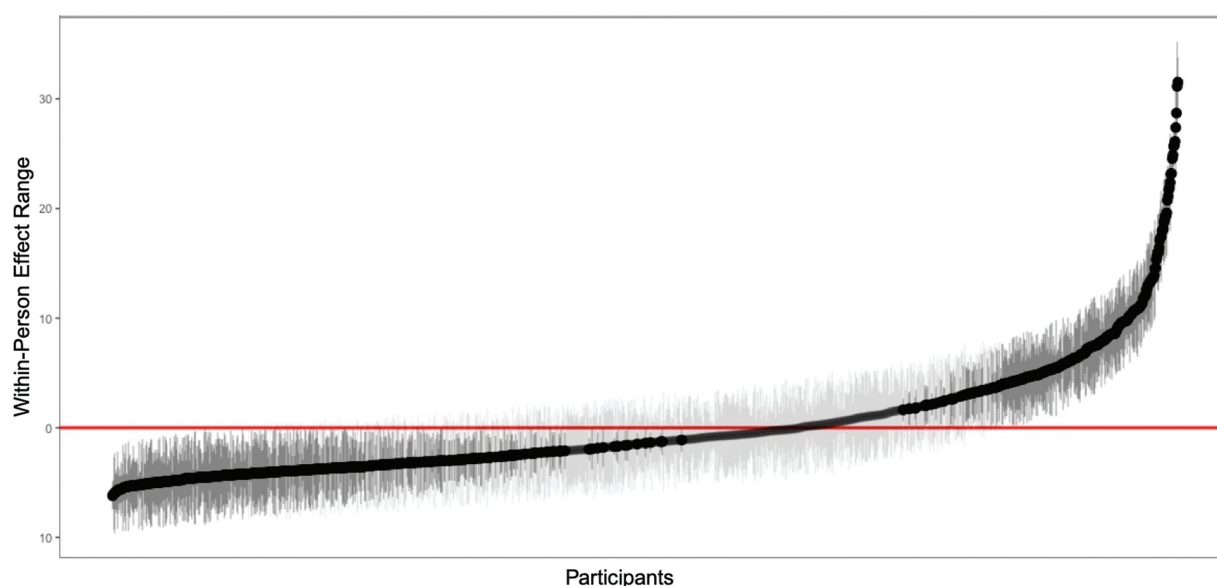
the time on average. Moreover, even though rest was the most prevalent technique, it was reported as helpful only 55.3% of the time on average. On the other hand, acupuncture and medical marijuana were the least prevalent (2.8 and 4.6%, respectively) but most frequently effective techniques on average (78.4 and 74.6% of the time, respectively).

## Discussion

In this study, we investigate the potential of using direct patient input via mobile technology to identify additional data elements that can be obtained for contributing to

existing data sources in the context of enigmatic disease: (1) improve prevalence estimates and characterization of traditionally underdocumented diseases, (2) provide multi-dimensional data (e.g., symptom severity, location, and temporal fluctuations<sup>84</sup>) on the disease for generating information that can aid in the personalization of care, and (3) provide contextual information to characterize burden of disease under real life circumstances and identify possible points of intervention. To our knowledge, this study is the first to provide a quantitative analysis to demonstrate how participatory research-based self-tracking data compare with EHR and what additional information could be extracted for investigation of underdocumented diseases.

Our finding that pelvic pain is the most prevalent symptom in our sample is in line with previous studies that use self-reported questionnaires to assess pain symptomology in endometriosis.<sup>45</sup> We extend this by reporting significantly fewer instances of pelvic pain and lack of any documentation of several other endometriosis-related pain problems (e.g., dyschezia<sup>85</sup>) in the EHR. Similarly, epigastric problems were significantly underdocumented in the EHR, even though this was the fourth most frequently reported problem in the Phendo sample. This could be due to various reasons. There could have been a higher number of patients presenting with epigastric pain, but these instances might have been documented as “abdominal pain” by the provider in the EHR without providing further details. From the patient’s



**Fig. 4** Plot of the person-level pain scores estimated from the multilevel model ( $n = 1,534$ ). Y-axis (“effect range”) represents pain scores. Each black dot represents one participant and gray lines indicate 95% confidence intervals. Distribution of points across the x-axis indicate large variability across individuals (i.e., between-group variance). Dark gray lines indicate random effects that are statistically significant.

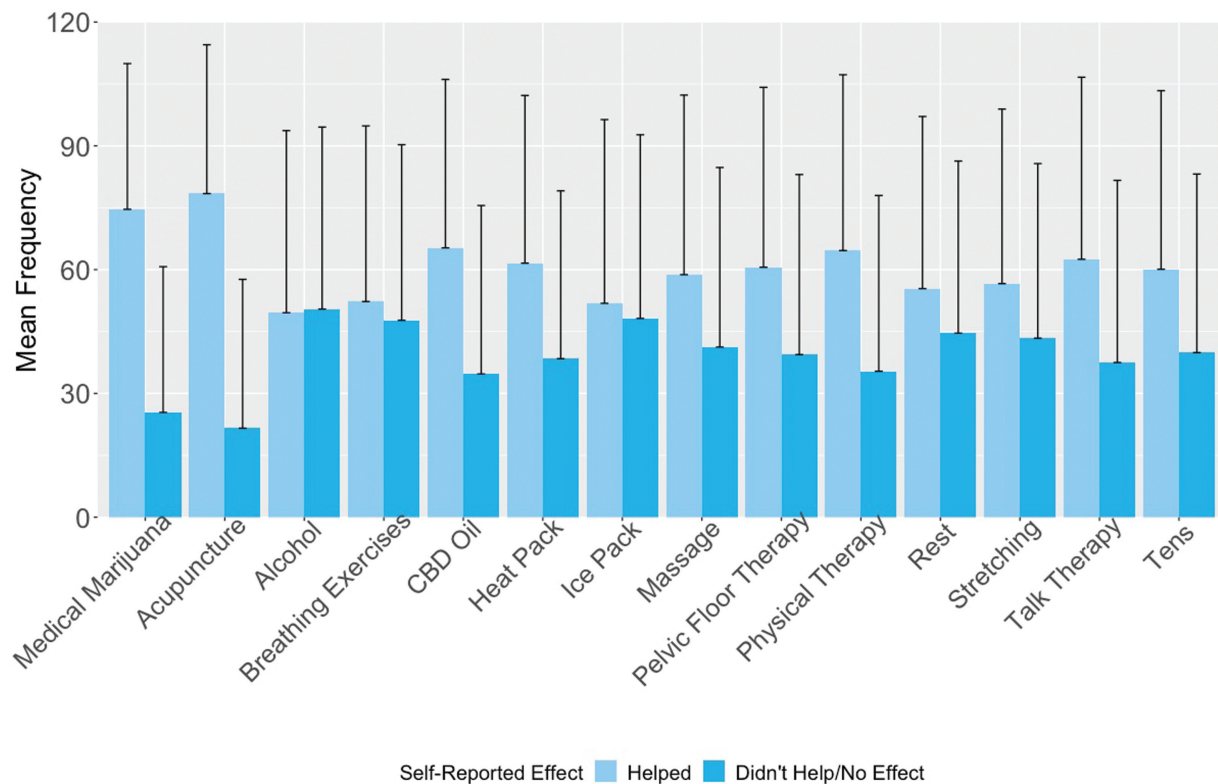
**Table 5** Prevalence of daily function tasks reported (%) in the Phendo sample

Daily function task	Count	Prevalence (%)
Working	3,681	53.1
Getting out of bed	3,619	52.2
Standing	3,287	47.4
Using the toilet	3,281	47.3
Sitting	3,273	47.2
Walking	3,214	46.4
Socializing	3,122	45.1
Sleeping	2,950	42.6
Dressing	2,706	39.0
Climbing stairs	2,334	33.7
Running	2,294	33.1
Eating	2,292	33.1
Preparing food	2,230	32.2
Stretching	2,191	31.6
Lying down	1,917	27.6
Jumping	1,812	26.1
Sex	1,778	25.6
Housework	1,755	25.3
Shopping	1,638	23.6
Lifting	1,469	21.2
Kneeling	1,309	18.9
Bathing	854	12.3

Note: Out of the total 6,925, all 5,764 participants tracked at least one daily function task in Phendo. Denominator of 6,925 is used to compute prevalence rates.

perspective, previous studies report patient resistance to disclosing pain,<sup>86</sup> and that patient’s physical pain can negatively affect patient–provider communication.<sup>87</sup> It is possible that patients feel less stigmatized when sharing their symptoms through a mobile app rather than directly sharing with their physician.<sup>53,88</sup> Another possibility could be lack of adequate health insurance for access to point of care, which has been linked to not only delays in diagnosis, but also disparities in primary and/or specialist care and treatment.<sup>40,89</sup> A similar study<sup>90</sup> comparing patient self-reported surveys to EHR problem lists and chart notes for common medical conditions reported that fewer than half of the 380 problems were shared across the data sources, and that 22% of all diagnoses were found only in patient self-report. Our findings reinforce these findings in demonstrating direct patient input provides an opportunity to bring these different data sources together to improve our understanding of patient health history and disease status.<sup>90</sup> We further demonstrate the added value of mHealth tools designed based on participatory research to promote patient engagement and self-tracking for improving our understanding of diseases that are underdocumented and clinically not well understood. Within the LHS framework, these have been identified as opportunities for patient inclusion to democratize the health care system.<sup>57</sup>

Assessment of variation in pain indicated substantial between- and within-individual variability in our sample, which is in line with findings from other chronic pain samples reported in the literature.<sup>74</sup> It has been long alluded to that endometriosis is a heterogeneous disease that fluctuates in symptoms within individual over time, and that the magnitude of this fluctuation is moreover variable across individuals, yet this has not been quantified. Our results provide quantitative support for the dynamic nature of



**Fig. 5** Average effect of self-management techniques reported by the Phendo participants ( $n = 6,025$ ). One-sided error bars were used for visual display and indicate standard deviations. Proportions of each technique tracked as helpful or unhelpful (i.e., “did not help” or “no effect”) out of the total number of times tracked was averaged across all participants. There was significant variability as indicated by the standard deviations. For example, for the average user, acupuncture was reported to be helpful approximately 75% of the time, and unhelpful approximately 25% of the time.

endometriosis pain and quantifying this variability is a starting point for being able to derive clinically relevant associations. That a patient's pain symptoms over time fluctuate almost as much as the variability in overall pain experience observed across a sample of patients indicates, at minimum, that information gathered about a patient at single contact time point (e.g., doctor office visit) might not be representative of the person's overall disease experience as it unfolds over time. Similarly, attempts to characterize endometriosis through data collection designs relying solely on EHR problem lists might yield incomplete or biased results, findings supported by previous studies conducting in other conditions.<sup>14,16,90</sup> From a clinical standpoint, pain variability has been demonstrated to be predictive of treatment response and disease outcomes.<sup>79,91</sup> For example, a study<sup>79</sup> reported that higher within-person pain variability at baseline in a randomized controlled trial for a fibromyalgia pain drug was predictive of placebo response and nonspecific treatment effects. In another study with fibromyalgia patients, authors demonstrated that between- and within-person variability predicted disease symptomology clusters and help identify disease phenotypes. In sum, frequent monitoring can capture this dynamic nature of chronic pain, yielding a more comprehensive and accurate disease profile than what can be estimated through EHR which relies on contact incidence.

Our observations on prevalence of impaired daily function are in line with others who report substantial disease

impact on QoL in endometriosis.<sup>43,50,52</sup> For example, Phendo participants reported work as their most affected task and almost half of the sample report their sleep being affected by their endometriosis, findings previously reported in a survey with 107 Puerto Rican women (mean age = 34.5) with endometriosis.<sup>43</sup> In addition, our findings on the prevalence of socializing and walking as affected daily function tasks are closely aligned with those reported in another sample of 193 Puerto Rican women with endometriosis (45.0 vs. 48% and 46.4 vs. 41.4%, respectively).<sup>52</sup> With respect to self-management techniques, acupuncture was self-reported as effective most of the time (~87%), though only 2.8% of the Phendo participants reporting its use. Similarly, physical therapy was reported by only 4.6% of the participants. These low frequencies might indicate lack of access to such treatments, based on prior evidence that insurance coverage, costs and changes in insurance policy have been identified as significant barriers to receiving musculoskeletal (e.g., physical and pelvic) therapy and rehabilitation services.<sup>92</sup> The ability to identify these nuances in patient experiences with respect to their disease management is another strength of using direct patient input via mobile technology. Though beyond the scope of the aims of this study, between-individual variability in the effects of self-management techniques merits further investigation. Taken together with our observations on the self-management technique effect variability, these findings not only underscore the burden of endometriosis and its impact on daily living, but also the potential of using

direct input from the patient for identifying suitable targets for intervention and effective strategies for symptom management.

While mHealth and self-tracking for self-management and improving health care have been widely advocated both by the scientific community<sup>93,94</sup> and public health agencies (e.g., WHO),<sup>38</sup> their use to generate better evidence about chronic diseases is still rare.<sup>95,96</sup> We demonstrate that self-tracking data can be valuable particularly for conducting observational research on pain-related diseases that are poorly understood and/or studied. Another strength of our study is the participatory design aspect of Phendo, which ensures that the items included in the app are relevant to the patient population. This has been identified as a way to improve study retention and long-term adherence to tracking,<sup>97</sup> which is particularly important when the goal is to increase documentation base for an enigmatic disease.<sup>60</sup> Indeed, the prevalence rate of endometriosis cases in our EHR database was 0.4%, which is substantially lower than the reported 10% prevalence estimated in the population. This provides further rationale for focusing efforts to augment the typically relied upon EHR databases through direct patient input to promote patient engagement and inclusion in the data generation process, which have been recommended as approaches for accomplishing the patient-centered-care system promoted within the LHS framework.<sup>57</sup> Self-tracking facilitates documentation of not only severe pain, but also mild and moderate pain instances. This can reduce the likelihood of over-representing severe cases, a potential limitation of EHR.<sup>41</sup> Daily tracking allows for monitoring acute pain, which can have distinctive differences in physiology compared with chronic pain,<sup>98</sup> but still be associated to subsequent risk for developing chronic pain in clinical populations,<sup>99,100</sup> particularly in women.<sup>101</sup> As such, information on temporal fluctuations can improve our understanding of the dynamic course of the disease and subsequently making timely clinical decisions.

We note the potential of patient self-recording for improving the completeness of demographic data in the health-care systems. This paucity of data on common demographic factors (e.g., race, ethnicity and education level)<sup>102</sup> and the exclusive reliance on ICD-9 coding<sup>103</sup> in the EHR is a commonly reported issue in the literature.<sup>104</sup> In a recent study,<sup>105</sup> among the 2.4 million patients in New York Presbyterian Hospital healthcare system's EHR (used also for extracting the EHR sample in this study), race, or ethnicity was unknown for 57%, compared with 86% when patients directly recorded themselves. Accomplishing this is critical for supporting precision medicine initiatives and reducing health care disparities and inequities.

We note several limitations in our study. First, we had independent samples of women with endometriosis, and as such there might have been differences in characteristics of the two samples that influence the proportion estimates. Next, we were unable to conduct statistical tests of significance for pain intensity, frequency of their fluctuations, and self-management techniques and daily function tasks

reported, as these data were not provided in the EHR for our sample. Third, EHR data from the clinical warehouse were available from January 1996 onward; therefore, our search precludes identification of cases prior to that year. Demographic factors were substantially missing in the EHR, an observation that has been consistently reported by others in the literature.<sup>102,103,105</sup> Likewise, some of the demographic information was also missing in the Phendo sample. The study samples therefore might not be representative of endometriosis patients of some socioeconomic and ethnic/racial backgrounds. Though we are not aware of race/ethnicity as a predictive risk factor for endometriosis diagnosis, we cannot ascertain if these results would have been different with a more diverse sample. Approximately, 60% of the Phendo participants self-reported residing in the United States and approximately 30% reported residing in other countries, whereas EHR data were obtained from a single (though ethnoracially and linguistically diverse) metropolitan city and we do not have data on the patients' living or working environments. As such, there could be differences between the two samples in comparison with respect to environmental factors that could have influenced the outcomes. Similarly, we do not know the exact breakdown of specialist versus general medicine practitioners using the EHR system. Though beyond the scope of the present study, future studies could investigate pain problem documentation by practitioner specialty area.

Finally, we note the challenges of investigations in the EHR, a topic extensively discussed in the literature,<sup>106,107</sup> which might have influenced our results. The ETL process that converts the clinical data into the OMOP data tables is a dynamic process that improves over time, but is not completely error free. As such, the EHR search might have missed some pain-related conditions. A recent study of EHR problem list completeness and accuracy reports that providers struggle more with maintaining complete problem lists for patients whose diseases status are less severe and less symptomatic.<sup>16</sup> Accordingly, some of the patients identified in the EHR cohort might have reported additional pain-related findings, but these might have not been documented by the provider.

## Conclusion

Patient self-tracking has been linked to increased self-efficacy, patient adherence, engagement as advocates in their personalized care, and improved health outcomes.<sup>108–110</sup> This approach can enhance the research process by increasing inclusion of those prone to health inequities,<sup>111</sup> including disease-level inequities as demonstrated herein using endometriosis as a grounding example. Direct patient input via self-tracking can facilitate collecting and sharing information in a timely manner without increasing provider documentation responsibility and associated burden. In the same way, the EHR are both the source of information and the target of interventions; mHealth applications could play a critical role by promoting patient engagement in aspects of the health care system.



## Clinical Relevance Statement

Use of mHealth technology for obtaining direct patient input can supplement clinical documentation captured during point of care to more accurately and comprehensively evaluate patient health history and status. Patient inclusion in the data generation and sharing process can further alleviate burden due to EHR use, and enable patient as an active participant in their care. This approach can be applied to other conditions beyond endometriosis, especially those that are chronic and fluctuating in symptoms over time.

## Multiple Choice Questions

1. Which of the following is a primary endometriosis-related problem detected in the self-tracking sample in line with previous literature, but was found to be significantly missing in the electronic health record (EHR)?

- a. Hip pain
- b. Pelvic pain
- c. Breast pain
- d. Headache

**Correct Answer:** The correct answer is option b. Pelvic pain is a primary characteristic symptom in endometriosis. It was the most prevalent problem in the Phendo sample (57.3%, compared with 29.8% in the EHR), a finding in line with previous studies that assess pain symptomology in endometriosis.<sup>45</sup> Upon comparison, we further report significantly fewer instances of epigastric pain and dysuria in the EHR, as well as other endometriosis-related pain problems (e.g., dyschezia<sup>85</sup>).

2. Which contributory aspect of direct patient input via mobile technology is investigated in this study?

- a. Increase the number of patients we can add into the Learning Health System (LHS)
- b. Increase the mHealth use among stakeholders of LHS
- c. Expand upon the characterization of burden of disease in the EHR
- d. Increase documentation on diseases that are underdocumented

**Correct Answer:** The correct answer is option d. While there are numerous potential benefits of including direct patient input to complement information gathered from EHR for conducting research and informing medical decisions, this study focuses on how these data could help generate additional information on underdocumented diseases (e.g., endometriosis) that would not be possible to obtain from EHR.

3. In addition to clinical conditions, what other type of information has been indicated to be significantly underdocumented in the EHR?

- a. Patient preferences for treatment
- b. Diagnosis dates
- c. Demographic factors
- d. Date of documentation

**Correct Answer:** The correct answer is option c. Information on common demographic factors (e.g., race, ethnicity, and education level) were substantially missing in the EHR in our analyses, an observation that has been also reported by others in the literature. For example, a study<sup>105</sup> reported that among the 2.4 million patients in a large New York City university health care system's EHR, race, or ethnicity was unknown for 57%, compared with 86% when patients directly recorded themselves.

4. This study investigates endometriosis pain variability based on patient self-reported data collected frequently over time. What do the results of the linear mixed model (LMM) with respect to the estimated intercepts (→Figure 4) indicate?

- a. The ratio of within-participant to group-level pain variability is comparable.
- b. The mean day-level pain scores substantially vary across participants.
- c. There is significant intraclass correlation of pain scores over time.
- d. Within-participant variation in daily pain accounts for 58% of the total model variance.

**Correct Answer:** The correct answer is option b. The use of LMM with a random intercept for participant allows estimation of a separate person-level mean in daily pain scores, represented by the intercept. This random effect was statistically significant in the model, indicating that there is substantial variability across participants in their daily pain levels. This is depicted in →Figure 4 where variability across participants is apparent across the x-axis, while within-participant variability is apparent across the y-axis, represented by the variable 95% confidence intervals.

### Protection of Human and Animal Subjects

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. All procedures followed were in accordance with ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000. Informed consent was obtained from all patients for being included in the study.

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

### Funding

This study received financial support from Columbia University Data Science Institute Postdoctoral Fellowship, Endometriosis Foundation of America, National Science Foundation (grant number: 1344668), and from National Institutes of Health, U.S. National Library of Medicine (grant number: R01 LM013043).

**Conflict of Interest**

None declared.

**References**

- Institute of Medicine (US) Roundtable on Evidence-Based Medicine Olsen L, Aisner D, McGinnis JM, eds. *The Learning Health-care System: Workshop Summary*. Washington, DC: National Academies Press; 2007
- McGinnis JM, Powers B, Grossmann C. *Digital Infrastructure for the Learning Health System: The Foundation for Continuous Improvement in Health and Health Care: Workshop Series Summary*. Washington, DC: National Academies Press; 2011
- Friedman CP, Wong AK, Blumenthal D. Achieving a nationwide learning health system. *Sci Transl Med* 2010;2(57):57cm29
- Riley WT, Glasgow RE, Etheredge L, Abernethy AP. Rapid, responsive, relevant (R3) research: a call for a rapid learning health research enterprise. *Clin Transl Med* 2013;2(01):10
- Greene SM, Reid RJ, Larson EB. Implementing the learning health system: from concept to action. *Ann Intern Med* 2012;157(03):207–210
- Green LW. Making research relevant: if it is an evidence-based practice, where's the practice-based evidence? *Fam Pract* 2008;25(Suppl 1):i20–i24
- Cortez A, Hsui P, Mitchell E, Riehl V, Smith P. Conceptualizing a data infrastructure for the capture, use, and sharing of patient-generated health data in care delivery and research through 2024. Accenture. . Available at: [https://www.healthit.gov/sites/default/files/onc\\_pghd\\_practical\\_guide.pdf?platform=hootsuite](https://www.healthit.gov/sites/default/files/onc_pghd_practical_guide.pdf?platform=hootsuite). Accessed January, 2018
- Casey JA, Schwartz BS, Stewart WF, Adler NE. Using electronic health records for population health research: a review of methods and applications. *Annu Rev Public Health* 2016;37:61–81
- Lin KJ, Singer DE, Glynn RJ, Murphy SN, Lii J, Schneeweiss S. Identifying patients with high data completeness to improve validity of comparative effectiveness research in electronic health records data. *Clin Pharmacol Ther* 2018;103(05):899–905
- Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care. *Nat Rev Genet* 2012;13(06):395–405
- Acker B, Bronnert J, Brown T, et al. Problem list guidance in the EHR. *J AHIMA* 2011;82(09):52–58
- Blondeau C. *Pocket Glossary of Health Information Management and Technology*. 3rd Edition Chicago, IL: AHIMA; 2011
- Daskivich TJ, Abedi G, Kaplan SH, et al. Electronic health record problem lists: accurate enough for risk adjustment? *Am J Manag Care* 2018;24(01):e24–e29
- Wang M, Cyhaniuk A, Cooper DL, Iyer NN. Identification of patients with congenital hemophilia in a large electronic health record database. *J Blood Med* 2017;8:131–139
- Wang YC, Shimbo D, Muntner P, Moran AE, Krakoff LR, Schwartz JE. Prevalence of masked hypertension among US adults with nonelevated clinic blood pressure. *Am J Epidemiol* 2017;185(03):194–202
- Wang EC-H, Wright A. Characterizing outpatient problem list completeness and duplications in the electronic health record. *J Am Med Inform Assoc* 2020;27(08):1190–1197
- Ommaya AK, Cipriano PF, Hoyt DB, et al. Care-centered clinical documentation in the digital environment: solutions to alleviate burnout. *NAM Perspectives* . Available at: <https://nam.edu/wp-content/uploads/2018/01/Care-Centered-Clinical-Documentation.pdf>. Accessed January 29, 2018
- Collier R. Electronic health records contributing to physician burnout. *Can Med Assoc* . Available at: <https://europepmc.org/article/med/29133547>. Accessed 2017
- Gardner RL, Cooper E, Haskell J, et al. Physician stress and burnout: the impact of health information technology. *J Am Med Inform Assoc* 2019;26(02):106–114
- West SL, Johnson W, Visscher W, Kluckman M, Qin Y, Larsen A. The challenges of linking health insurer claims with electronic medical records. *Health Informatics J* 2014;20(01):22–34
- Adane K, Gizachew M, Kendie S. The role of medical data in efficient patient care delivery: a review. *Risk Manag Healthc Policy* 2019;12:67–73
- Rhodes ET, Laffel LM, Gonzalez TV, Ludwig DS. Accuracy of administrative coding for type 2 diabetes in children, adolescents, and young adults. *Diabetes Care* 2007;30(01):141–143
- Valikodath NG, Newman-Casey PA, Lee PP, Musch DC, Niziol LM, Woodward MA. Agreement of ocular symptom reporting between patient-reported outcomes and medical records. *JAMA Ophthalmol* 2017;135(03):225–231
- Cahill S, Makadon H. Sexual orientation and gender identity data collection in clinical settings and in electronic health records: a key to ending LGBT health disparities. *LGBT Health* 2014;1(01):34–41
- Bagley SC, Altman RB. Computing disease incidence, prevalence and comorbidity from electronic medical records. *J Biomed Inform* 2016;63:108–111
- Almalki M, Gray K, Sanchez FM. The use of self-quantification systems for personal health information: big data management activities and prospects. *Health Inf Sci Syst* 2015;3(Suppl 1):S1
- Swan M. Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified self-tracking. *Int J Environ Res Public Health* 2009;6(02):492–525
- Khovanskaya V, Baumer EPS, Cosley D, Volda S, Gay G. Everybody knows what you're doing: a critical design approach to personal informatics. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* . Available at: <https://stephen.volda.com/uploads/Publications/Publications/khovanskaya-chi13.pdf>. Accessed 2013
- Figueiredo M, Caldeira C, Chen Y, Zheng K. Routine self-tracking of health: reasons, facilitating factors, and the potential impact on health management practices. *AMIA Annu Symp Proc* 2018;2017:706–714
- Shapiro M, Johnston D, Wald J, Mon D. Patient-generated health data. RTI International . Available at: [https://www.healthit.gov/sites/default/files/rti\\_pghd\\_whitepaper\\_april\\_2012.pdf](https://www.healthit.gov/sites/default/files/rti_pghd_whitepaper_april_2012.pdf). Accessed April 2012
- Boyer GS, Templin DW, Goring WP, et al. Discrepancies between patient recall and the medical record. Potential impact on diagnosis and clinical assessment of chronic disease. *Arch Intern Med* 1995;155(17):1868–1872
- Sampson UKA, Kaplan RM, Cooper RS, et al. Reducing health inequities in the U.S.: recommendations from the NHLBI's health inequities think tank meeting. *J Am Coll Cardiol* 2016;68(05):517–524
- Ahmed S, Berzon RA, Revicki DA International Society for Quality of Life Research. , et al; . The use of patient-reported outcomes (PRO) within comparative effectiveness research: implications for clinical practice and health care policy. *Med Care* 2012;50(12):1060–1070
- Pichon A, Schiffer K, Horran E, Bakken S, Mamykina L, Elhadad N. Divided We Stand: The Collaborative Work of Patients and Providers in an Enigmatic Chronic Disease. [under anonymous review]. . Available at: [http://www.columbia.edu/~ab3886/docs/APichon\\_CV\\_Jan2020.pdf](http://www.columbia.edu/~ab3886/docs/APichon_CV_Jan2020.pdf). Accessed 2020
- Payne TH, Corley S, Cullen TA, et al. Report of the AMIA EHR-2020 task force on the status and future direction of EHRs. *J Am Med Inform Assoc* 2015;22(05):1102–1110
- Rycroft-Malone J, Seers K, Titcher A, Harvey G, Kitson A, McCormack B. What counts as evidence in evidence-based practice? *J Adv Nurs* 2004;47(01):81–90

- 37 Dullabh P, Hovey L, Heaney-Huls K, Rajendran N, Wright A, Sittig DF. Application programming interfaces in health care: findings from a current-state sociotechnical assessment. *Appl Clin Inform* 2020;11(01):59–69
- 38 Executive B. mHealth: use of appropriate digital technologies for public health: report by the Director-General. Geneva World Health Organization; Available at: [https://apps.who.int/iris/bitstream/handle/10665/274134/B142\\_20-en.pdf?sequence=1&isAllowed=y](https://apps.who.int/iris/bitstream/handle/10665/274134/B142_20-en.pdf?sequence=1&isAllowed=y). Accessed November 27, 2017
- 39 Tamareis JS, Irwin JC, Goldfien GA, et al. Molecular classification of endometriosis and disease stage using high-dimensional genomic data. *Endocrinology* 2014;155(12):4986–4999
- 40 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;221(06):623.e1–623.e11
- 41 De Graaff AA, D'Hooghe TM, Dunselman GA, Dirksen CD, Hummelshoj L, Simoens SWERF EndoCost Consortium. The significant effect of endometriosis on physical, mental and social wellbeing: results from an international cross-sectional survey. *Hum Reprod* 2013;28(10):2677–2685
- 42 Simoens S, Dunselman G, Dirksen C, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. *Hum Reprod* 2012;27(05):1292–1299
- 43 Fourquet J, Gao X, Zavala D, et al. Patients' report on how endometriosis affects health, work, and daily life. *Fertil Steril* 2010;93(07):2424–2428
- 44 Rogers PA, D'Hooghe TM, Fazleabas A, et al. Priorities for endometriosis research: recommendations from an international consensus workshop. *Reprod Sci* 2009;16(04):335–346
- 45 Schliep KC, Mumford SL, Peterson CM, et al. Pain typology and incident endometriosis. *Hum Reprod* 2015;30(10):2427–2438
- 46 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. *J Manag Care Spec Pharm* 2017;23(07):745–754
- 47 Whiteman MK, Hillis SD, Jamieson DJ, et al. Inpatient hysterectomy surveillance in the United States, 2000–2004. *Am J Obstet Gynecol* 2008;198(01):34.e1–34.e7
- 48 Health Nio Estimates of funding for various research, condition, and disease categories (RCDC). Available at: [https://report.nih.gov/categorical\\_spending.aspx](https://report.nih.gov/categorical_spending.aspx). Published 2019. Accessed March 15, 2020
- 49 Garry R. Is insulin resistance an essential component of PCOS?: the endometriosis syndromes: a clinical classification in the presence of aetiological confusion and therapeutic anarchy. *Hum Reprod* 2004;19(04):760–768
- 50 Nnoaham KE, Hummelshoj L, Webster P World Endometriosis Research Foundation Global Study of Women's Health consortium. , et al; . Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries. *Fertil Steril* 2011;96(02):366–373
- 51 Farland LV, Horne AW. Disparity in endometriosis diagnoses between racial/ethnic groups. *BJOG* 2019;126(09):1115–1116
- 52 Fourquet J, Báez L, Figueroa M, Iriarte RI, Flores I. Quantification of the impact of endometriosis symptoms on health-related quality of life and work productivity. *Fertil Steril* 2011;96(01):107–112
- 53 Chen EH, Shofer FS, Dean AJ, et al. Gender disparity in analgesic treatment of emergency department patients with acute abdominal pain. *Acad Emerg Med* 2008;15(05):414–418
- 54 Leresche L. Defining gender disparities in pain management. *Clin Orthop Relat Res* 2011;469(07):1871–1877
- 55 Ghiasi M, Kulkarni MT, Missmer SA. Is endometriosis more common and more severe than it was 30 years ago? *J Minim Invasive Gynecol* 2020;27(02):452–461
- 56 Budrionis A, Bellika JG. The learning healthcare system: where are we now? A systematic review. *J Biomed Inform* 2016;64:87–92
- 57 Galvin HK, Petersen C, Subbian V, Solomonides A. Patients as agents in behavioral health research and service provision: recommendations to support the learning health system. *Appl Clin Inform* 2019;10(05):841–848
- 58 McKillop M, Mamykina L, Elhadad N. Designing in the dark: eliciting self-tracking dimensions for understanding enigmatic disease. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems . Available at: <http://people.dbmi.columbia.edu/noemie/papers/18chi.pdf>. Accessed 2018
- 59 McKillop M, Voigt N, Schnall R, Elhadad N. Exploring self-tracking as a participatory research activity among women with endometriosis. *J Particip Med* 2016;8:e17
- 60 Ensari I, Elhadad N. mHealth For Research: Participatory Research Applications to Gain Disease Insights. In *Digital Health: Mobile and Wearable Devices for Participatory Health Applications*. Eds. Syed-Abdul S, Zhu X, Fernandez-Luque L. Elsevier, 2020. Forthcoming. ISBN: 978-0128200773
- 61 Friedman C, Rubin J, Brown J, et al. Toward a science of learning systems: a research agenda for the high-functioning Learning Health System. *J Am Med Inform Assoc* 2015;22(01):43–50
- 62 Hripcsak G, Duke JD, Shah NH, et al. Observational health data sciences and informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform* 2015;216:574–578
- 63 Hripcsak G, Albers DJ. Correlating electronic health record concepts with healthcare process events. *J Am Med Inform Assoc* 2013;20(e2):e311–e318
- 64 United States Census Bureau QuickFacts. New York City, New York United States Census Bureau. Available at: <https://data.census.gov/cedsci/all?q=new%20york%20city&g=1600000US3651000&hidePreview=false&tid=ACSDP1-Y2018.DP05&vintage=2018>. Accessed 2020
- 65 Ta CN, Dumontier M, Hripcsak G, Tatonetti NP, Weng C. Columbia Open Health Data, clinical concept prevalence and co-occurrence from electronic health records. *Sci Data* 2018;5:180273–180273
- 66 Awad E, Ahmed HAH, Yousef A, Abbas R. Efficacy of exercise on pelvic pain and posture associated with endometriosis: within subject design. *J Phys Ther Sci* 2017;29(12):2112–2115
- 67 The Observational Medical Outcomes Partnership (OMOP) Common Data Model. Available at: <https://ohdsi.github.io/Common-DataModel/background.html>. Accessed 2020
- 68 International Health Terminology Standards Development Organization SNOMED CT Worldwide. Available at: <http://www.ihtsdo.org/snomed-ct/snomed-ct-worldwide>. Accessed 2016
- 69 Zini EM, Lanzola G, Bossi P, Quaglini S. An environment for guideline-based decision support systems for outpatients monitoring. *Methods Inf Med* 2017;56(04):283–293
- 70 Ramakrishnan N, Hanauer D, Keller B. Mining electronic health records. *Computer* 2010;(10):77–81
- 71 McKillop MM. Phenotyping endometriosis from observational health data. Columbia University . Available at: <https://academiccommons.columbia.edu/doi/10.7916/d8-1est-dh56>. Accessed March 18, 2019
- 72 Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975;1(03):277–299
- 73 Simborg DW, Starfield BH, Horn SD, Yourtee SA. Information factors affecting problem follow-up in ambulatory care. *Med Care* 1976;14(10):848–856
- 74 Bartley EJ, Robinson ME, Staud R. Pain and fatigue variability patterns distinguish subgroups of fibromyalgia patients. *J Pain* 2018;19(04):372–381
- 75 Salthouse TA. Implications of within-person variability in cognitive and neuropsychological functioning for the interpretation of change. *Neuropsychology* 2007;21(04):401–411
- 76 Urteaga I, McKillop M, Elhadad N. Learning endometriosis phenotypes from patient-generated data. *NPJ Digit Med* 2020;3:88

- 77 Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc* 1927;22(158):209–212
- 78 Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. *Stat Med* 1998;17(08):873–890
- 79 Harris RE, Williams DA, McLean SA, et al. Characterization and consequences of pain variability in individuals with fibromyalgia. *Arthritis Rheum* 2005;52(11):3670–3674
- 80 Stoffel MA, Nakagawa S, Schielzeth H. rptR: Repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods Ecol Evol* 2017;8(11):1639–1644
- 81 Hox JJ, Moerbeek M, Van de Schoot R. *Multilevel Analysis: Techniques and Applications*. Routledge 2010
- 82 Nakagawa S, Schielzeth H. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biol Rev Camb Philos Soc* 2010;85(04):935–956
- 83 Huang C-W, Lu R, Iqbal U, et al. A richly interactive exploratory data analysis and visualization tool using electronic medical records. *BMC Med Inform Decis Mak* 2015;15:92–92
- 84 Bougie O, Healey J, Singh SS. Behind the times: revisiting endometriosis and race. *Am J Obstet Gynecol* 2019;221(01):35.e1–35.e5
- 85 Seracchioli R, Mabrouk M, Guerrini M, et al. Dyschezia and posterior deep infiltrating endometriosis: analysis of 360 cases. *J Minim Invasive Gynecol* 2008;15(06):695–699
- 86 Cagle J, Bunting M. Patient reluctance to discuss pain: understanding stoicism, stigma, and other contributing factors. *J Soc Work End Life Palliat Care* 2017;13(01):27–43
- 87 Norouzinia R, Aghabarari M, Shiri M, Karimi M, Samami E, Samami EJGjohs. Communication barriers perceived by nurses and patients. *Glob J Health Sci* 2015;8(06):65–74
- 88 Hoffmann DE, Tarzian AJ. The girl who cried pain: a bias against women in the treatment of pain. *J Law Med Ethics* 2001;29(01):13–27
- 89 As-Sanie S, Black R, Giudice LC, et al. Assessing research gaps and unmet needs in endometriosis. *Am J Obstet Gynecol* 2019;221(02):86–94
- 90 Weiskopf NG, Cohen AM, Hannan J, Jarmon T, Dorr DA. Towards augmenting structured EHR data: a comparison of manual chart review and patient self-report. *AMIA Annu Symp Proc* 2020;2019:903–912
- 91 Farrar JT, Troxel AB, Haynes K, et al. Effect of variability in the 7-day baseline pain diary on the assay sensitivity of neuropathic pain randomized clinical trials: an ACTION study. *Pain* 2014;155(08):1622–1631
- 92 Carvalho E, Bettger JP, Goode AP. Insurance coverage, costs, and barriers to care for outpatient musculoskeletal therapy and rehabilitation services. *N C Med J* 2017;78(05):312–314
- 93 Genes N, Violante S, Cetrangol C, Rogers L, Schadt EE, Chan YY. From smartphone to EHR: a case report on integrating patient-generated health data. *NPJ Digit Med* 2018;1(01):23
- 94 Alexander S. mHealth Technologies for the Self-management of Diabetes in the Older Population. *ACM SIGACCESS Accessibility and Computing* 2015;(111):14–18
- 95 Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. *Sci Transl Med* 2015;7(283):283rv3
- 96 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(01):7
- 97 Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med* 2017;7(02):254–267
- 98 Honoré P, Menning PM, Rogers SD, Nichols ML, Mantyh PW. Neurochemical plasticity in persistent inflammatory pain. In: . *Progress in brain research*. Vol 129 Elsevier; 2000:357–363
- 99 Sinatra R. Causes and consequences of inadequate management of acute pain. *Pain Med* 2010;11(12):1859–1871
- 100 Tasmuth T, Kataja M, Blomqvist C, von Smitten K, Kalso E. Treatment-related factors predisposing to chronic pain in patients with breast cancer—a multivariate approach. *Acta Oncol* 1997;36(06):625–630
- 101 Poleshuck EL, Katz J, Andrus CH, et al. Risk factors for chronic pain following breast cancer surgery: a prospective study. *J Pain* 2006;7(09):626–634
- 102 Fremont A, Lurie N. Eliminating health disparities: measurement and data needs. Appendix D, The role of racial and ethnic data collection in eliminating disparities in health care Washington (DC): National Academies Press (US); 2004
- 103 Birkhead GS, Klompas M, Shah NR. Uses of electronic health records for public health surveillance to advance public health. *Annu Rev Public Health* 2015;36(01):345–359
- 104 Douglas MD, Dawes DE, Holden KB, Mack D. Missed policy opportunities to advance health equity by recording demographic data in electronic health records. *Am J Public Health* 2015;105(Suppl 3):S380–S388
- 105 Polubriaginof FCG, Ryan P, Salmasian H, et al. Challenges with quality of race and ethnicity data in observational databases. *J Am Med Inform Assoc* 2019;26(8-9):730–736
- 106 Kiepek W, Sengstack PP. An evaluation of system end-user support during implementation of an electronic health record using the model for improvement framework. *Appl Clin Inform* 2019;10(05):964–971
- 107 Thyvalikakath TP, Duncan WD, Siddiqui ZNational Dental PBRN Collaborative Group. , et al; . Leveraging electronic dental record data for clinical research in the national dental PBRN practices. *Appl Clin Inform* 2020;11(02):305–314
- 108 Charnel PA, Frampton SB. Building the business case for patient-centered care. *Healthc Financ Manage* 2008;62(03):80–85
- 109 Chiauzzi E, Rodarte C, DasMahapatra P. Patient-centered activity monitoring in the self-management of chronic health conditions. *BMC Med* 2015;13(01):77
- 110 Korotitsch WJ, Nelson-Gray RO. An overview of self-monitoring research in assessment and treatment. *Psychol Assess* 1999;11(04):415
- 111 Chase HS, Radhakrishnan J, Shirazian S, Rao MK, Vawdrey DK. Under-documentation of chronic kidney disease in the electronic health record in outpatients. *J Am Med Inform Assoc* 2010;17(05):588–594