Original Paper

Exploring Self-Reported Symptoms for Developing and Evaluating Digital Symptom Checkers for Polycystic Ovarian Syndrome, Endometriosis, and Uterine Fibroids: Exploratory Survey Study

Aidan P Wickham¹, PhD; Yella Hewings-Martin¹, PhD; Frederick GB Goddard¹, PhD; Allison K Rodgers², MD; Adam C Cunningham¹, PhD; Carley Prentice¹, MPH; Octavia Wilks¹, MBChB; Yusuf C Kaplan¹, MD; Andrei Marhol¹, PhD, MD; András Meczner¹, MD; Heorhi Stsefanovich¹, BSc; Anna Klepchukova¹, MD; Liudmila Zhaunova¹, PhD

¹Flo Health UK Limited, London, United Kingdom
 ²Fertility Centers of Illinois, Vernon Hills, IL, United States

Corresponding Author:

Aidan P Wickham, PhD Flo Health UK Limited 27 Old Gloucester Street London, WC1N 3AX United Kingdom Phone: 44 44 60396823 Email: <u>a_wickham@flo.health</u>

Abstract

Background: Reproductive health conditions such as polycystic ovary syndrome (PCOS), endometriosis, and uterine fibroids pose a significant burden to people who menstruate, health care systems, and economies. Despite clinical guidelines for each condition, prolonged delays in diagnosis are commonplace, resulting in an increase to health care costs and risk of health complications. Symptom checker apps have the potential to significantly reduce time to diagnosis by providing users with health information and tools to better understand their symptoms.

Objective: This study aims to study the prevalence and predictive importance of self-reported symptoms of PCOS, endometriosis, and uterine fibroids, and to explore the efficacy of 3 symptom checkers (developed by Flo Health UK Limited) that use self-reported symptoms when screening for each condition.

Methods: Flo's symptom checkers were transcribed into separate web-based surveys for PCOS, endometriosis, and uterine fibroids, asking respondents their diagnostic history for each condition. Participants were aged 18 years or older, female, and living in the United States. Participants either had a confirmed diagnosis (condition-positive) and reported symptoms retrospectively as experienced at the time of diagnosis, or they had not been examined for the condition (condition-negative) and reported their current symptoms as experienced at the time of surveying. Symptom prevalence was calculated for each condition based on the surveys. Least absolute shrinkage and selection operator regression was used to identify key symptoms for predicting each condition. Participants' symptoms were processed by Flo's 3 single-condition symptom checkers, and accuracy was assessed by comparing the symptom checker output with the participant's condition designation.

Results: A total of 1317 participants were included with 418, 476, and 423 in the PCOS, endometriosis, and uterine fibroids groups, respectively. The most prevalent symptoms for PCOS were fatigue (92%), feeling anxious (87%), BMI over 25 (84%); for endometriosis: very regular lower abdominal pain (89%), fatigue (85%), and referred lower back pain (80%); for uterine fibroids: fatigue (76%), bloating (69%), and changing sanitary protection often (68%). Symptoms of anovulation and amenorrhea (long periods, irregular cycles, and absent periods), and hyperandrogenism (excess hair on chin and abdomen, scalp hair loss, and BMI over 25) were identified as the most predictive symptoms for PCOS, while symptoms related to abdominal pain and the effect pain has on life, bleeding, and fertility complications were among the most predictive symptoms for both endometriosis and uterine fibroids. Symptom checker accuracy was 78%, 73%, and 75% for PCOS, endometriosis, and uterine fibroids, respectively.

Conclusions: This exploratory study characterizes self-reported symptomatology and identifies the key predictive symptoms for 3 reproductive conditions. The Flo symptom checkers were evaluated using real, self-reported symptoms and demonstrated high levels of accuracy.

RenderX

Wickham et al

KEYWORDS

polycystic ovary syndrome; PCOS; self-assessment; self-reported; endometriosis; uterine fibroids; symptoms; digital symptom checker; women's health; gynecological conditions; reproductive health

Introduction

Background

Polycystic ovary syndrome (PCOS), endometriosis, and uterine fibroids are among the most prevalent reproductive health conditions, collectively impacting millions of women and people who menstruate globally. Prevalences of PCOS and endometriosis range from 4%-20% and 5%-10%, respectively [1-4]. For uterine fibroids, prevalence of diagnosis is approximately 10% across the female population [5], and importantly, the condition has a high estimated cumulative incidence by onset of menopause, occurring in over 70% of women [6,7].

Despite the symptoms of these conditions being well-established in medical guidelines, diagnostic delays are common, with patients enduring between 2 and 12 years from the onset of their symptoms before they receive a formal diagnosis [8-10]. These delays could be due to complexities in diagnostic procedure [11], controversy in diagnostic evaluation [12], or affected persons having low knowledge of reproductive conditions and their symptoms [13].

Self-assessed, self-reported symptoms are known to be an effective means of identifying those who need medical attention [14], and their severity has been shown to correlate with diagnosis [15,16]. Better understanding of how those affected by these reproductive conditions self-assess and self-report their symptoms is vital, as left untreated these conditions can lead to worsening of symptoms, infertility, and further health complications [13,17-19]. People are increasingly trusting mobile apps and the internet as a source of health information [20-22] and intervention to assist with explaining or recognizing their symptoms [23,24]. "Symptom checkers" are a digital health tool that allow users to match their symptoms against those of health conditions. Symptom checkers cannot provide diagnosis to the user, but they do hold the potential to facilitate screening of reproductive conditions at scale, and therefore, rigorous validation and proof of their accuracy is essential [25]. A validation process that uses self-reported symptoms from real people will help to understand how a symptom checker will perform once deployed.

Flo App and Symptom Checker Development

Flo (by Flo Health UK Limited) is a period-tracking mobile app [26] allowing users to log their menstruation, ovulation, contraception, symptoms, and other physical and emotion-based metrics. Flo has developed 3 single-condition symptom checkers for PCOS, endometriosis, and uterine fibroids [27] that are based on current guidelines (Monash, AAFP, and ESHRE) [28-30]. Interactions with these are either user-led or triggered by symptoms or events logged in the app (eg, detection of cycle irregularity). The symptom checkers ask users conversation-like questions to gather information on symptoms, medical, and

https://formative.jmir.org/2024/1/e65469

XSL•FO

family history related to each condition. The symptom checkers are not diagnostic tools; the user receives an outcome of "significant match," "low match," or "no match," and an informative summary of their symptomatology that can be used to facilitate conversations with their health care provider should they wish to seek subsequent medical investigations. The symptom checkers use a cumulative total of present symptoms over a threshold to determine the condition-match output the user receives.

Aim

This exploratory study sought to estimate self-reported symptomatology of PCOS, endometriosis, and uterine fibroids, and investigate which symptoms were most predictive of each condition. We also explored the potential accuracy of three single-condition symptom checkers when assessing self-reported symptoms.

Methods

Self-Reported Symptom Data Collection

Participants

Subscribers to the paid survey platform "Prolific" were invited to take part in a study related to female reproductive health from September to December 2023. Participants were included if they consented, were aged 18 years or older, female, and living in the United States. Due to the relatively low prevalence of each condition, participants were prescreened to ensure there were enough participants with a condition diagnosis included in the study. Participants were invited to complete a survey related to either PCOS, endometriosis, or uterine fibroids. Informed electronic consent was obtained from all participants at the start of the survey.

Surveys

The questions from each single-condition symptom checker were transcribed into 3 separate surveys. The surveys comprised the following sections: (1) examination history for the condition in question, (2) medical history relevant to the condition, and (3) the questions from each symptom checker. In addition, the PCOS symptom checker contains images which show examples of hyperandrogenism symptoms, which was reflected in the PCOS survey.

Respondents who indicated they had a positive diagnosis for PCOS, endometriosis, or uterine fibroids were asked to answer symptom questions retrospectively. Respondents who did not have a diagnosis for one of these 3 conditions answered the surveys about their current symptoms.

Case Designation

Condition-positive (CP) cases were designated as those participants who had received a positive diagnosis for PCOS,

endometriosis, or uterine fibroids from their doctor. It is possible CP participants were multimorbid with 2 or more of the conditions investigated in this study.

Condition-negative (CN) cases were designated as those participants who responded they had not been examined by a doctor for the condition. We did not use participants with a confirmed negative diagnosis (ie, examined for the condition and told they do not have it) as our CN cases, as their visits to a doctor might indicate potential symptoms for other conditions, making them unsuitable as control cases. We assumed that most undiagnosed, unexamined participants will be negative for the condition; it is possible that CN cases may unknowingly have the condition, but this bias is expected to be relatively small due to the symptomatic prevalence of each condition within reproductive age being approximately 10% or less [1,3,5,31].

Analysis

Participant Refinement

CP participants were excluded from the analysis if they were outside the ages of 18 and 45 years at the time of their diagnosis (calculated from their age and year of diagnosis, provided in their survey responses), while CN participants were excluded if they were outside the ages of 18 and 45 years at the time of survey completion. As CN participants answered questions about their current symptom status (as opposed to their retrospective symptom status), CN participants were excluded if they were currently pregnant, had recently given birth, or were currently taking hormonal contraceptives, in case these elements interfered with their symptomatology. PCOS-negative participants were also excluded if they had recently had a miscarriage due to the effect this can have on menstrual cycle regularity. Endometriosis-negative and uterine fibroids-negative participants were not excluded on this criterion, as miscarriage has been shown to be associated with both conditions [32-34].

By definition, our symptom checkers can only detect symptomatic cases of each condition. Due to the high-rate of uterine fibroids cases being asymptomatic [35], we asked uterine fibroids-positive participants if they were symptomatic or not at the time of their diagnosis. Participants who responded "I was asymptomatic (had no symptoms)" were excluded from the analysis. Participants with PCOS and endometriosis were not asked if they were asymptomatic at the time of their diagnosis, and so our analysis may include some asymptomatic cases of these 2 conditions. Participants who did not complete the entire survey were removed from the analysis.

Symptom Presence

The presence or absence of a symptom was recorded for each participant based on their answer to each survey question. Neutral responses to questions such as "I don't know" or "I'd prefer not to answer" were recorded as an absent symptom.

Statistical Analysis

 $XSI \bullet FC$

Symptom prevalence was calculated relative to the number of individuals in each CP or CN group; for symptoms related to fertility or pregnancy, symptom prevalence was calculated relative to those participants who had indicated they either had previously tried or were currently trying to conceive. Chi-square

```
https://formative.jmir.org/2024/1/e65469
```

test without correction was used to test statistical difference in symptom prevalence between CP and CN groups.

To select the most important symptoms for detecting a CP outcome from our self-reported symptom dataset, we used the use of least absolute shrinkage and selection operator (LASSO) regression. The LASSO regression provided feature importance scores for the most predictive symptoms of the CP outcome, from which we present the top 10 symptoms for each condition.

Participants' symptoms were processed by each symptom checker's cumulative algorithm and the outcome was recorded as either a "significant match (symptoms suggestive of the condition) or a "low or no match" (symptoms not suggestive enough of the condition). To produce accuracy metrics, "significant match" was considered equivalent to CP, and "low or no match" was considered equivalent to CN. Accuracy was calculated as the sum of correctly identified CP and CN cases divided by the total number of cases for each condition.

Sensitivity Analysis

A sensitivity analysis investigated self-reported symptom prevalence in participants with a confirmed-negative diagnosis (examined for the condition and told they do not have it). We also tested the performance of the Flo symptom checkers using this confirmed-negative group.

Ethical Considerations

This study and protocol were approved by the Independent Ethical Review Board: WIRB-Copernicus Group Institutional Review Board (IRB number 20231821). All participants completed an electronic, informed consent form, agreeing to their health and personal data being processed. Participants were informed their responses would be deidentified and all published results would be aggregated. Participants were compensated approximately \$11.62/hour (£9/hour) for completing the survey.

Results

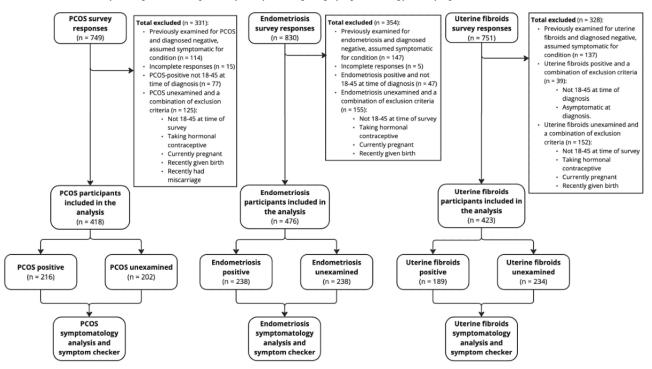
Sample Characteristics

A total of 2330 survey responses were collected, with 749, 830, and 751 responses collected for PCOS, endometriosis, and uterine fibroids, respectively (Figure 1). After exclusion, we collected 216 and 238 responses that had a positive diagnosis for PCOS and endometriosis, and 189 responses that had a positive, symptomatic diagnosis for uterine fibroids that we determined as our CP cases. From CN cases, those who had not been examined for the condition, we included responses from 202 participants for PCOS, 238 for endometriosis, and 234 for uterine fibroids.

The CP mean age at diagnosis was 26.6 (SD 5.7), 28.6 (SD 6.3), and 33.4 (SD 6.8) years for PCOS, endometriosis, and uterine fibroids, respectively (Table S1 in Multimedia Appendix 1). The CP group for uterine fibroids had the highest share of Black people (19.6%), compared with PCOS (8.8%) and endometriosis (8.0%; Table S1 in Multimedia Appendix 1). Among CP groups, the rate of condition-family history was similar for PCOS and endometriosis (28.7% and 33.6%, respectively) and higher for the uterine fibroids group (45.0%). Among CN groups, the rate

of condition-family history for PCOS, endometriosis, and uterine fibroids was similar at 10.9%, 10.1%, and 12%, respectively.

Figure 1. Participants flow chart. A total of 749 polycystic ovary syndrome (PCOS), 830 endometriosis, and 751 uterine fibroids symptom survey responses were collected from September to December, 2023 and filtered according to the exclusion criteria. In total, 216 PCOS-positive and 202 PCOS unexamined, 238 endometriosis-positive and 238 endometrioses unexamined, and 189 uterine fibroids-positive and 234 uterine fibroids unexamined were taken forward for analysis as part of an exploratory study investigating symptomatology and symptom checker evaluation.



Self-Reported Symptomatology

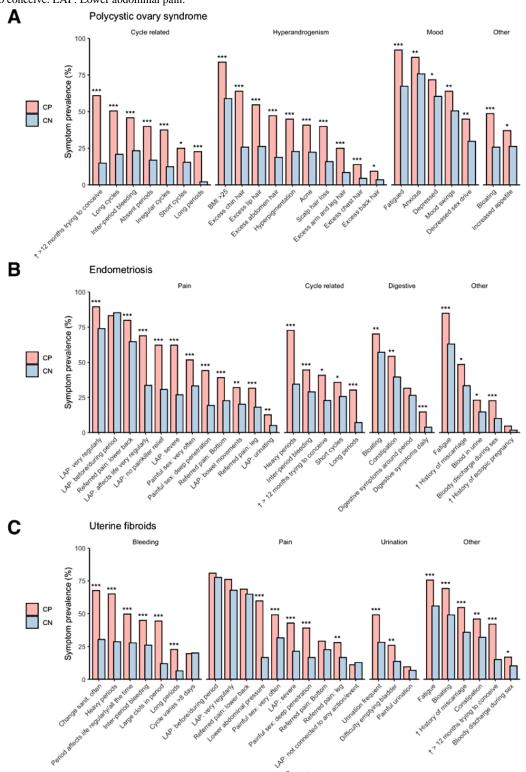
All included symptoms were significantly more prevalent among PCOS-positive compared with PCOS-negative participants (Figure 2A and Table S2A in Multimedia Appendix 1) with "Fatigue" (92%) being the most prevalent. "Excess chin hair" was reported considerably more by the PCOS-positive group (64%) than within the PCOS-negative group (26%). "BMI over 25" was highly prevalent in the PCOS-positive group (84%) and was also reported among the majority (59%) in the PCOS-negative group. PCOS-positive participants reported absent periods (amenorrhea) and irregular cycles 40% and 38% of the time, respectively.

The most prevalent symptoms reported by endometriosis-positive participants, with a significant difference to endometriosis-negative participants, were "Lower abdominal pain: very regularly" (89%), "Fatigue" (85%), and "Referred pain: lower back" (80%; Figure 2B and Table S3B in Multimedia Appendix 1). "Lower abdominal pain before or during period" (dysmenorrhea) was highly prevalent in both the endometriosis-positive and negative groups (83% and 85%) with no significant difference between them (P=.53). Pain that affects the participant's life regularly ("Lower abdominal pain: affects life very regularly") was much more prevalent in the endometriosis-positive group compared with the negative group (69% and 34%).

The most prevalent symptoms reported by uterine fibroids-positive participants, with a significant difference to uterine fibroids-negative respondents, were "Fatigue" (76%), "Bloating" (69%), and "Changing sanitary protection often" (68%; Figure 2C and Table S3C in Multimedia Appendix 1). "Lower abdominal pain before or during period" was the most prevalent symptom in both the uterine fibroids-positive and negative groups (81% and 73%) with no statistical significance between them (P=.42). "Lower abdominal pressure" was the most prevalent pain symptom amongst uterine fibroids-positive participants (60%) with a large difference compared with the negative group (17%).



Figure 2. Self-reported symptomatology for condition-positive and condition-negative participants for (A) polycystic ovary syndrome, (B) endometriosis, and (C) uterine fibroids. Symptoms were self-reported and in response to survey questions from Flo single-condition symptom checkers. Statistical significance between groups is indicated: *P<.05; **P<.01; ***P<.001. \dagger symptom prevalence is reported for those who have previously tried or are currently trying to conceive. LAP: Lower abdominal pain.



Symptoms

Amongst participants with a current or previous history of trying to conceive, 60% of PCOS-positive participants, 41% of endometriosis-positive participants, and 42% of uterine fibroids-positive participants reported they were trying or had tried to conceive for 12 months or longer compared with 15%, 23%, and 15% of PCOS-, endometriosis- and uterine fibroids-negative groups, respectively. Amongst the same group

https://formative.jmir.org/2024/1/e65469

XSL•FO RenderX of participants, a history of miscarriage was reported at 48% and 55% for endometriosis- and uterine fibroids-positive participants compared with 33% and 36% for endometriosis- and uterine fibroids-negative participants.

A sensitivity analysis investigating symptomatology of participants with a confirmed negative diagnosis (after

examination) for each condition found this group had a higher prevalence of nearly all symptoms compared with the CN group used in our main analysis (Figure S1 in Multimedia Appendix 1).

Symptom Importance

The 10 most predictive symptoms of each condition identified by LASSO regression are presented in Table 1. Interestingly

for PCOS, certain locations of hirsutism (abdomen and chin) were identified as important, indicating the self-reporting of other hirsutism locations were not as predictive of the condition. For endometriosis and uterine fibroids, many shared symptoms (ie, painful sex, heavy periods, and complications with fertility) have been identified as important for condition prediction.

Table 1. Top 10 most predictive symptoms for identifying polycystic ovary syndrome, endometriosis, and uterine fibroids, determined by least absolute shrinkage and selection operator regression. Symptoms were self-reported and in response to survey questions from Flo single-condition symptom checkers from condition-positive and condition-negative participants.

PCOS ^a	Endometriosis	Uterine fibroids
Cycle		
Long periods	b	_
Absent periods	_	_
Cycle varies 8 days	varies 8 days —	
Interperiod bleeding	_	_
Pain		
_	Severe lower abdominal pain	Lower abdominal pressure
_	Painful sex: deep penetration	Painful sex: deep penetration
_	Pain affects life regularly/always	_
_	No relief from painkillers	_
Bleeding		
_	Heavy periods	Heavy periods
_	Long periods	Long periods
_	_	Change tampon/pad regularly
_	_	Period clots
Hyperandrogenism		
Excess chin hair	_	_
Excess abdominal hair	_	_
Scalp hair loss	_	_
BMI>25	_	_
Fertility		
_	Trying to conceive >12 months	Trying to conceive >12 months
_	History of miscarriage	History of miscarriage
Bladder control		
_	_	Frequent urination
_	_	Difficulty emptying bladder
Other		
Fatigue	Fatigue	_
Bloating	Daily digestive symptoms	_

^aPCOS: polycystic ovary syndrome.

^b—: not applicable.

RenderX

Symptom Checker Accuracy

For symptom checker performance, we found an accuracy of 78% for PCOS, 73% for endometriosis, and 75% for uterine fibroids (Table 2). Sensitivity and specificity were 76% and 81% for PCOS, 73% and 73% for endometriosis, and 78% and 73% for uterine fibroids.

The sensitivity analysis considering participants with a confirmed negative diagnosis for the condition found that accuracy reduced slightly due to a decrease in specificity (Table S3 in Multimedia Appendix 1).

Table 2. Performance metrics for polycystic ovary syndrome, endometriosis, and uterine fibroids symptom checkers. Symptoms used in symptom checker evaluation were self-reported and in responses to survey questions containing the same questions as the Flo symptom checkers. Symptom checker performance was evaluated when classifying between condition-positive and condition-negative cases as either "significant match" or "low or no match."

Metric	PCOS ^a	Endometriosis	Uterine fibroids
CP ^b , n	216	238	189
CN ^c , n	202	238	234
Accuracy (%)	78	73	75
Sensitivity (%)	76	73	78
Specificity (%)	81	73	73
PPV ^d (%)	81	73	70
NPV ^e (%)	76	73	81

^aPCOS: polycystic ovary syndrome.

^bCP: condition positive.

^cCN: condition negative.

^dPPV: positive predictive value.

^eNPV: negative predictive value.

Discussion

Preliminary Findings

In this exploratory study, we created and distributed surveys based on Flo's symptom checkers for PCOS, endometriosis, and uterine fibroids. We collected self-reported symptoms, analyzed their prevalence and predictive value for each condition, and evaluated the symptom checker's performance. For each condition, the top 3 most prevalent symptoms with a statistically significant difference between the CP and CN group were as follows: for PCOS-positive, "Fatigue" (92%, P<.001), "Anxious" (87%, P=.003), and "BMI over 25" (84%, P<.001), for endometriosis-positive, "Lower abdominal pain: very regularly" (89%, P<.001), "Fatigue" (85%, P<.001), and "Referred pain: lower back" (80%, P<.001), "Bloated" (69%, P<.001), "Change sanitary products often" (68%, P<.001).

There was no statistical difference in prevalence for "Lower abdominal pain: before or during period," "Lower abdominal pain: very regularly," and "Referred pain: lower back" between CP and CN groups for endometriosis and uterine fibroids participants, despite these symptoms being commonly associated with both conditions. For some symptoms, such as "Fatigue," "Anxious," "Lower abdominal pain: very regularly," despite the difference in prevalence between CP and CN groups being statistically significant, the prevalence in the CN group was still high. These findings suggest that some symptoms that are clinically recognized as being associated with these conditions

RenderX

in medical guidelines are commonly self-reported by healthy people as well.

For PCOS, LASSO regression identified particular symptoms related to cycle irregularity (varied cycle length, long periods, and absent periods) and hyperandrogenism as being most important for condition prediction. Interestingly, current guidelines for PCOS state that the presence of scalp hair loss and acne in isolation (without hirsutism) are weak indicators of biochemical hyperandrogenism [29], yet our analysis finds them among the most important symptoms in predicting PCOS, possibly indicating there are differences in importance between self-reported symptoms and clinical evaluation.

Endometriosis and uterine fibroids have multiple symptoms that overlap between the 2 conditions [36]. LASSO regressions identified the same symptoms related to heavy, long periods, fertility, and pain with deep penetration were identified as being important for condition prediction. The self-reported symptoms which the 2 conditions do not have in common, but were identified as predictive (ie, severe, life affecting abdominal pain for endometriosis compared with a feeling of constant abdominal pressure for uterine fibroids) may support clinical decision making and are described as such in the different guidelines. Focusing on the difference in these pain-related symptoms could give clinicians more confidence when differentiating between them. Fatigue was identified as one of the most important symptoms for the prediction of PCOS and endometriosis despite being highly prevalent in the CN groups.

We tested single-condition symptom checkers using self-reported symptom datasets and found accuracy ranged from 73%-78%. Each symptom checker performed with high sensitivity (73%-78%) and specificity (73%-81%), successfully identifying people who do and do not have the target condition to a high degree. This shows that the symptom checkers can screen people for PCOS, endometriosis, and uterine fibroids when using self-assessed, self-reported symptoms with a high level of accuracy.

Comparison to Previous Work

Symptom-based questionnaires and screening tools do exist for reproductive health conditions such as PCOS and endometriosis. A self-assessment tool for endometriosis found sensitivity of 76% and specificity of 72%, however, the accuracy metrics are based on a relatively small number of participants (n=50) compared with our study [37]. Another study that used a questionnaire for use in PCOS diagnosis has reported 77% sensitivity and 94% specificity. However, the study surveyed women who had an existing primary complaint of infertility [38]. It is possible that this could bias performance as not all people with PCOS will experience infertility [39]. The Flo symptom checkers are designed for a broader audience who may not be experiencing fertility issues or unaware that they could have a health condition.

We previously tested the performance of Flo's symptom checkers using vignettes (simulated patient cases) [27]. We reported the accuracy in our vignettes study to be higher (83%-88%) compared with this survey-based study. This was expected, as vignette cases were created from a panel of clinical experience, portraying classic presentations of each condition. Self-reported patient data, as collected in our study, reflect a real-world scenario with greater nuances in symptomatology on which to test our symptom checkers.

Strengths and Limitations

Strengths of this study include the large number of CP and CN participants used to evaluate each symptom checker, likely covering extensive variations of symptomatology. This has an advantage over vignette-case studies where simulated patient cases are unlikely to represent the complexity of real patient symptomatology [40-42]. The participants used in the main analysis of this study may have multiple reproductive or other health conditions with symptom crossover, providing a complex, realistic dataset of self-reported symptoms to test Flo's symptom checkers. Each condition survey contained questions transcribed from as they appear in Flo's symptom checker, and their completion of surveys was completely participant led with no clinical guidance or coaching to aid answering questions or identify symptoms. Thus, the accuracy metrics for each symptom checker are likely indicative of how the Flo symptom checkers would perform once deployed.

Limitations, however, should be noted. We asked CP participants to answer questions in the survey retrospectively, to the time of their diagnosis. It is possible that self-reported symptoms from CP participants could contain bias or inaccuracies, particularly in participants who were diagnosed many years ago [40,43]. Recall bias can affect those who have

been diagnosed or previously examined, resulting in participants who are more likely to report symptoms [44]. Our sensitivity analysis showed participants with a confirmed negative diagnosis generally self-reported a higher symptom prevalence for all symptoms compared with the CN group in our main analysis. This could be due to recall bias, however another explanation is that these participants sought medical examination for these conditions due to the symptoms they were experiencing. Another limitation of this study is that the CP and CN groups could have been matched more closely by age and ethnicity. These conditions can vary in prevalences across different ages and ethnicity groups, as well as the presentation of symptoms.

The CN participants used in this study had not been examined for the condition by a doctor, and it was our assumption that the large majority of this group are negative for the condition due to the relatively low symptomatic prevalence of each condition. This assumption has been used by other studies in their assessment of machine-learning tools for endometriosis prediction [45,46]. However, due to this imperfect reference category, it is likely there is a bias in the metrics reported. When refining our uterine fibroids participants, we excluded participants who said they were asymptomatic at the time of their diagnosis. The intention of this approach was to exclude participants who have asymptomatic uterine fibroids; however, it may have also removed participants who were unaware their symptoms were related to uterine fibroids.

Finally, the Flo symptom checker questions, from which we estimated the prevalence of self-reported symptomatology of each condition are based on current medical guidelines. As a result, it is possible that the symptomatology and importance of symptoms presented in this study could be biased by diagnostic constructs and entrenched clinical unknowns for each condition. Further work could investigate a larger array of self-reported general and reproductive health symptoms that extend beyond what symptoms are currently clinically accepted for each condition.

Future Work

The evaluation of all digital health tools should follow a multistage process with increasing exposure to real environments [25]. The next step in evaluating Flo symptom checkers is early-field testing; using real-world data collected from users using the symptom checker, and comparing the symptom checker's output to the user's official diagnosis they received from their doctor. The symptom checker performance will be continually reviewed and adjusted using the most up to date evidence and guidelines as part of an iterative product development process to ensure its safety and accuracy.

Symptom checkers that provide accurate screening of reproductive conditions could help facilitate conversations with health care providers and assist in clinical evaluation, reducing the time to diagnosis and decreasing the risk of complications caused from untreated conditions that drive up health care costs [47-49]. Access to health tools, such as symptom checkers, for conditions like PCOS, endometriosis, and uterine fibroids holds the potential to be a cost-effective strategy for reducing this economic burden, especially when evaluation represents a

XSL•FO

fraction of the all-cost total [50], and warrants further study. Future work should investigate the potential economic impact of the symptom checkers and the cost-saving potentials to health care systems and economies.

Conclusions

Characterization of reproductive conditions by self-reported symptoms could lead to more effective screening and is vital when developing digital health solutions such as symptom

Acknowledgments

checkers. In this exploratory study, we have presented the self-reported symptomatology for PCOS, endometriosis, and uterine fibroids, and analyzed the importance of each self-reported symptom when identifying each condition. Under these testing conditions, the Flo symptom checkers show high levels of accuracy, sensitivity, and specificity when assessing real, self-reported symptoms. Innovative health app solutions hold the potential to help people seek treatment before further morbidities and health complications occur.

This study was funded by Flo Health UK Limited. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

Due to the sensitivity of the data subject (ie, reproductive health symptom data), we have restricted the availability of the data. The data used for this study can be made available in aggregated format through reasonable requests to the corresponding author.

Conflicts of Interest

APW, YHM, ACC, OW, YCK, A Marhol, A Meczner, HS, and AK, are employees at Flo Health and hold equity interest in Flo Health. FG and CP are consultants for Flo Health. LZ is a consultant for Flo Health and holds equity interest. AKR is a consultant for Flo Health and a member of the medical advisory board

Multimedia Appendix 1

Supplementary materials. [DOCX File , 433 KB-Multimedia Appendix 1]

References

- 1. Azziz R, Carmina E, Chen Z, Dunaif A, Laven JSE, Legro RS, et al. Polycystic ovary syndrome. Nat Rev Dis Primers. 2016;2:16057. [doi: 10.1038/nrdp.2016.57] [Medline: 27510637]
- 2. Deswal R, Narwal V, Dang A, Pundir CS. The prevalence of polycystic ovary syndrome: a brief systematic review. J Hum Reprod Sci. 2020;13(4):261-271. [FREE Full text] [doi: 10.4103/jhrs.JHRS_95_18] [Medline: 33627974]
- 3. Taylor HS, Kotlyar AM, Flores VA. Endometriosis is a chronic systemic disease: clinical challenges and novel innovations. Lancet. 2021;397(10276):839-852. [doi: 10.1016/S0140-6736(21)00389-5] [Medline: 33640070]
- Rowlands I, Abbott J, Montgomery G, Hockey R, Rogers P, Mishra G. Prevalence and incidence of endometriosis in Australian women: a data linkage cohort study. BJOG. 2021;128(4):657-665. [FREE Full text] [doi: 10.1111/1471-0528.16447] [Medline: 32757329]
- 5. Zimmermann A, Bernuit D, Gerlinger C, Schaefers M, Geppert K. Prevalence, symptoms and management of uterine fibroids: an international internet-based survey of 21,746 women. BMC Womens Health. 2012;12(1):6. [FREE Full text] [doi: 10.1186/1472-6874-12-6] [Medline: 22448610]
- 6. Stewart E, Cookson C, Gandolfo R, Schulze-Rath R. Epidemiology of uterine fibroids: a systematic review. BJOG. 2017;124(10):1501-1512. [doi: 10.1111/1471-0528.14640] [Medline: 28296146]
- Baird DD, Dunson DB, Hill MC, Cousins D, Schectman JM. High cumulative incidence of uterine leiomyoma in black and white women: ultrasound evidence. Am J Obstet Gynecol. 2003;188(1):100-107. [doi: <u>10.1067/mob.2003.99</u>] [Medline: <u>12548202</u>]
- Agarwal SK, Chapron C, Giudice LC, Laufer MR, Leyland N, Missmer SA, et al. Clinical diagnosis of endometriosis: a call to action. Am J Obstet Gynecol. 2019;220(4):354.e1-354.e12. [FREE Full text] [doi: 10.1016/j.ajog.2018.12.039] [Medline: 30625295]
- 9. Borah BJ, Nicholson WK, Bradley L, Stewart EA. The impact of uterine leiomyomas: a national survey of affected women. Am J Obstet Gynecol. 2013;209(4):319.e1-319.e20. [doi: <u>10.1016/j.ajog.2013.07.017</u>] [Medline: <u>23891629</u>]
- Gibson-Helm ME, Lucas IM, Boyle JA, Teede HJ. Women's experiences of polycystic ovary syndrome diagnosis. Fam Pract. 2014;31(5):545-549. [doi: <u>10.1093/fampra/cmu028</u>] [Medline: <u>24925927</u>]
- 11. Spaczynski RZ, Duleba AJ. Diagnosis of endometriosis. Semin Reprod Med. 2003;21(2):193-208. [doi: 10.1055/s-2003-41326] [Medline: 12917789]

RenderX

- Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Hum Reprod. 2018;33(9):1602-1618. [FREE Full text] [doi: 10.1093/humrep/dey256] [Medline: 30052961]
- Ghant MS, Sengoba KS, Vogelzang R, Lawson AK, Marsh EE. An altered perception of normal: understanding causes for treatment delay in women with symptomatic uterine fibroids. J Womens Health (Larchmt). 2016;25(8):846-852. [doi: <u>10.1089/jwh.2015.5531</u>] [Medline: <u>27195902</u>]
- Bastiani L, Fortunato L, Pieroni S, Bianchi F, Adorni F, Prinelli F, et al. Rapid COVID-19 screening based on self-reported symptoms: psychometric assessment and validation of the EPICOVID19 short diagnostic scale. J Med Internet Res. 2021;23(1):e23897. [FREE Full text] [doi: 10.2196/23897] [Medline: 33320825]
- 15. Rutter LA, Howard J, Lakhan P, Valdez D, Bollen J, Lorenzo-Luaces L. "I Haven't been diagnosed, But I Should Be"-insight into self-diagnoses of common mental health disorders: cross-sectional study. JMIR Form Res. 2023;7:e39206. [FREE Full text] [doi: 10.2196/39206] [Medline: 36637885]
- 16. Santos BF, Oliveira HN, Miranda AES, Hermsdorff HHM, Bressan J, Vieira JCM, et al. Research quality assessment: Reliability and validation of the self-reported diagnosis of depression for participants of the Cohort of Universities of Minas Gerais (CUME project). Journal of Affective Disorders Reports. 2021;6:100238. [doi: 10.1016/j.jadr.2021.100238]
- Lizneva D, Suturina L, Walker W, Brakta S, Gavrilova-Jordan L, Azziz R. Criteria, prevalence, and phenotypes of polycystic ovary syndrome. Fertil Steril. 2016;106(1):6-15. [FREE Full text] [doi: 10.1016/j.fertnstert.2016.05.003] [Medline: 27233760]
- 18. Witchel SF, Teede HJ, Peña AS. Curtailing PCOS. Pediatr Res. 2020;87(2):353-361. [doi: <u>10.1038/s41390-019-0615-1</u>] [Medline: <u>31627209</u>]
- 19. Ghai V, Jan H, Shakir F, Haines P, Kent A. Diagnostic delay for superficial and deep endometriosis in the United Kingdom. J Obstet Gynaecol. 2020;40(1):83-89. [FREE Full text] [doi: 10.1080/01443615.2019.1603217] [Medline: 31328629]
- 20. Fox S, Duggan M. Health Online 2013. Pew research center; 2013. URL: <u>https://www.pewresearch.org/internet/2013/01/</u> 15/health-online-2013/ [accessed 2024-11-27]
- 21. Yigzaw KY, Wynn R, Marco-Ruiz L, Budrionis A, Oyeyemi SO, Fagerlund AJ, et al. The association between health information seeking on the internet and physician visits (The Seventh Tromsø Study Part 4): Population-Based Questionnaire Study. J Med Internet Res. 2020;22(3):e13120. [FREE Full text] [doi: 10.2196/13120] [Medline: 32134387]
- Meyer AND, Giardina TD, Spitzmueller C, Shahid U, Scott TMT, Singh H. Patient perspectives on the usefulness of an artificial intelligence-assisted symptom checker: cross-sectional survey study. J Med Internet Res. 2020;22(1):e14679. [FREE Full text] [doi: 10.2196/14679] [Medline: 32012052]
- 23. Jutel A, Lupton D. Digitizing diagnosis: a review of mobile applications in the diagnostic process. Diagnosis De Gruyte. 2015;2(2):89-96. [doi: 10.1515/dx-2014-0068] [Medline: 29540025]
- 24. Wetzel AJ, Koch R, Preiser C, Müller R, Klemmt M, Ranisch R, et al. Ethical, legal, and social implications of symptom checker apps in primary health care (CHECK.APP): protocol for an interdisciplinary mixed methods study. JMIR Res Protoc. 2022;11(5):e34026. [FREE Full text] [doi: 10.2196/34026] [Medline: 35576570]
- 25. Fraser H, Coiera E, Wong D. Safety of patient-facing digital symptom checkers. Lancet. 2018;392(10161):2263-2264. [doi: 10.1016/S0140-6736(18)32819-8] [Medline: 30413281]
- 26. We're Flo, the world's #1 women's health app. Flo Health URL: <u>https://flo.health/</u> [accessed 2024-11-27]
- Peven K, Wickham AP, Wilks O, Kaplan YC, Marhol A, Ahmed S, et al. Assessment of a digital symptom checker tool's accuracy in suggesting reproductive health conditions: clinical vignettes study. JMIR Mhealth Uhealth. 2023;11:e46718.
 [FREE Full text] [doi: 10.2196/46718] [Medline: 38051574]
- Becker CM, Bokor A, Heikinheimo O, Horne A, Jansen F, Kiesel L, et al. ESHRE guideline: endometriosis. Hum Reprod Open. 2022;2022(2):hoac009. [FREE Full text] [doi: 10.1093/hropen/hoac009] [Medline: 35350465]
- 29. Teede H, Tay CT, Laven JSE, Dokras A, Moran LJ, Piltonen T, et al. Recommendations from the 2023 international evidence-based guideline for the assessment and management of polycystic ovary syndrome. J Clin Endocrinol Metab. 2023;108(10):2447-2469. [doi: 10.1210/clinem/dgad463] [Medline: 37580314]
- De La Cruz MSD, Buchanan EM. Uterine fibroids: diagnosis and treatment. Am Fam Physician. 2017;95(2):100-107.
 [FREE Full text] [Medline: 28084714]
- Giuliani E, As-Sanie S, Marsh EE. Epidemiology and management of uterine fibroids. Int J Gynaecol Obstet. 2020;149(1):3-9.
 [FREE Full text] [doi: 10.1002/ijgo.13102] [Medline: 31960950]
- Kohl Schwartz AS, Wölfler MM, Mitter V, Rauchfuss M, Haeberlin F, Eberhard M, et al. Endometriosis, especially mild disease: a risk factor for miscarriages. Fertil Steril. 2017;108(5):806-814. [FREE Full text] [doi: 10.1016/j.fertnstert.2017.08.025] [Medline: 29079275]
- 33. Santulli P, Marcellin L, Menard S, Thubert T, Khoshnood B, Gayet V, et al. Increased rate of spontaneous miscarriages in endometriosis-affected women. Hum Reprod. 2016;31(5):1014-1023. [doi: <u>10.1093/humrep/dew035</u>] [Medline: <u>26965434</u>]
- Saravelos SH, Yan J, Rehmani H, Li TC. The prevalence and impact of fibroids and their treatment on the outcome of pregnancy in women with recurrent miscarriage. Hum Reprod. 2011;26(12):3274-3279. [doi: <u>10.1093/humrep/der293</u>] [Medline: <u>21954281</u>]
- 35. Divakar H. Asymptomatic uterine fibroids. Best Pract Res Clin Obstet Gynaecol. 2008;22(4):643-654. [doi: 10.1016/j.bpobgyn.2008.01.007] [Medline: 18375184]



- 36. Uimari O, Nazri H, Tapmeier T. Endometriosis and uterine fibroids (Leiomyomata): comorbidity, risks and implications. Front Reprod Health. 2021;3:750018. [FREE Full text] [doi: 10.3389/frph.2021.750018] [Medline: 36304022]
- Cho HH, Yoon YS. Development of an endometriosis self-assessment tool for patient. Obstet Gynecol Sci. 2022;65(3):256-265. [FREE Full text] [doi: 10.5468/ogs.21252] [Medline: 35381626]
- 38. Pedersen SD, Brar S, Faris P, Corenblum B. Polycystic ovary syndrome: validated questionnaire for use in diagnosis. Can Fam Physician. 2007;53(6):1042-1047. [FREE Full text] [Medline: <u>17872783</u>]
- 39. Joham AE, Norman RJ, Stener-Victorin E, Legro RS, Franks S, Moran LJ, et al. Polycystic ovary syndrome. The Lancet Diabetes & Endocrinology. 2022;10(9):668-680. [doi: <u>10.1016/s2213-8587(22)00163-2</u>]
- 40. Wallace W, Chan C, Chidambaram S, Hanna L, Iqbal FM, Acharya A, et al. The diagnostic and triage accuracy of digital and online symptom checker tools: a systematic review. NPJ Digit Med. 2022;5(1):118. [FREE Full text] [doi: 10.1038/s41746-022-00667-w] [Medline: 35977992]
- 41. El-Osta A, Webber I, Alaa A, Bagkeris E, Mian S, Sharabiani MTA, et al. What is the suitability of clinical vignettes in benchmarking the performance of online symptom checkers? An audit study. BMJ Open. 2022;12(4):e053566. [FREE Full text] [doi: 10.1136/bmjopen-2021-053566] [Medline: 35477872]
- 42. Williams B, Song JJY. Are simulated patients effective in facilitating development of clinical competence for healthcare students? A scoping review. Adv Simul (Lond). 2016;1(1):6. [FREE Full text] [doi: 10.1186/s41077-016-0006-1] [Medline: 29449975]
- 43. Van den Bergh O, Walentynowicz M. Accuracy and bias in retrospective symptom reporting. Curr Opin Psychiatry. 2016;29(5):302-308. [doi: 10.1097/YCO.00000000000267] [Medline: 27427854]
- 44. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. J Multidiscip Healthc. 2016;9:211-217. [FREE Full text] [doi: 10.2147/JMDH.S104807] [Medline: 27217764]
- 45. Goldstein A, Cohen S. Self-report symptom-based endometriosis prediction using machine learning. Sci Rep. 2023;13(1):5499. [FREE Full text] [doi: 10.1038/s41598-023-32761-8] [Medline: 37016132]
- 46. Bendifallah S, Puchar A, Suisse S, Delbos L, Poilblanc M, Descamps P, et al. Machine learning algorithms as new screening approach for patients with endometriosis. Sci Rep. 2022;12(1):639. [doi: <u>10.1038/s41598-021-04637-2</u>] [Medline: <u>35022502</u>]
- 47. Ballweg ML. Impact of endometriosis on women's health: comparative historical data show that the earlier the onset, the more severe the disease. Best Pract Res Clin Obstet Gynaecol. 2004;18(2):201-218. [doi: 10.1016/j.bpobgyn.2004.01.003] [Medline: 15157638]
- 48. Peña AS, Witchel SF, Hoeger KM, Oberfield SE, Vogiatzi MG, Misso M, et al. Adolescent polycystic ovary syndrome according to the international evidence-based guideline. BMC Med. 2020;18(1):72. [doi: <u>10.1186/s12916-020-01516-x</u>] [Medline: <u>32204714</u>]
- Surrey E, Soliman AM, Trenz H, Blauer-Peterson C, Sluis A. Impact of endometriosis diagnostic delays on healthcare resource utilization and costs. Adv Ther. 2020;37(3):1087-1099. [FREE Full text] [doi: 10.1007/s12325-019-01215-x] [Medline: 31960340]
- 50. Azziz R. Overview of long-term morbidity economic cost of the polycystic ovary syndrome. In: Azziz R, Nestler JE, Dewailly D, editors. Androg Excess Disord Women Polycystic Ovary Syndr Disord. Totowa. Humana Press; 2007.

Abbreviations

CN: condition-negativeCP: condition-positiveLASSO: Least absolute shrinkage and selection operatorPCOS: polycystic ovary syndrome

Edited by A Mavragani; submitted 16.08.24; peer-reviewed by J Doust, R Marshall; comments to author 05.10.24; revised version received 25.10.24; accepted 13.11.24; published 12.12.24

Please cite as:

Exploring Self-Reported Symptoms for Developing and Evaluating Digital Symptom Checkers for Polycystic Ovarian Syndrome, Endometriosis, and Uterine Fibroids: Exploratory Survey Study

JMIR Form Res 2024;8:e65469

URL: <u>https://formative.jmir.org/2024/1/e65469</u> doi: <u>10.2196/65469</u>

PMID:

Wickham AP, Hewings-Martin Y, Goddard FGB, Rodgers AK, Cunningham AC, Prentice C, Wilks O, Kaplan YC, Marhol A, Meczner A, Stsefanovich H, Klepchukova A, Zhaunova L

©Aidan P Wickham, Yella Hewings-Martin, Frederick GB Goddard, Allison K Rodgers, Adam C Cunningham, Carley Prentice, Octavia Wilks, Yusuf C Kaplan, Andrei Marhol, András Meczner, Heorhi Stsefanovich, Anna Klepchukova, Liudmila Zhaunova. Originally published in JMIR Formative Research (https://formative.jmir.org), 12.12.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative.jmir.org, as well as this copyright and license information must be included.