

Research and Applications

Can menstrual health apps selected based on users' needs change health-related factors? A double-blind randomized controlled trial

Jisan Lee¹ and Jeongeun Kim^{2,3,4}

¹Department of Nursing science, College of Life & Health Sciences, Hoseo University, Asan, Korea, ²Interdisciplinary Program of Medical Informatics, Seoul National University, Seoul, Korea, ³College of Nursing, Seoul National University, Seoul, Korea, and ⁴Research Institute of Nursing Science, Seoul National University, Seoul, Korea

Corresponding Author: Jeongeun Kim, PhD, RN, Informatics Nurse Specialist, College of Nursing Seoul National University, 103 Daehak-ro, Jongno-gu, Seoul 110-799, Korea (jeongeunkim0424@gmail.com)

Received 3 October 2018; Revised 6 January 2019; Editorial Decision 29 January 2019; Accepted 17 February 2019

ABSTRACT

Objective: Most healthcare providers are reluctant to use health apps for healthcare because there is no rigorous way of choosing the best app for their patient or consumer. Accordingly, we developed a new method of app selection that fully considers target users' needs. This study verified whether health apps selected based on target users' needs can influence health-related factors.

Materials and Methods: We conducted a randomized control trial of women with dysmenorrhea and premenstrual syndrome using App A (the best app selected using the new method) and App B (the app with the highest number of users worldwide). The intervention was performed over 4 months to include at least 3 menstrual cycles.

Results: Sixty-one app users completed the 16-week intervention. While users rated both apps as higher in quality than previously used menstrual apps, only App A users showed significant improvements in overall satisfaction, app outcome expectancy, the number of days with records, app social influence, intent to recommend, and the possibility of behavioral or cognitive changes in their symptom management. The number of menus used increased over time. While the app self-efficacy and the number of relief methods did not significantly differ between groups, they still showed an increase in App A users.

Conclusions: When a menstrual app reflected users' needs, they recorded their symptoms more often and reported higher app quality, satisfaction, and intention to recommend. This study can not only benefit the selection of menstrual apps, but also confirm that mobile health apps can improve health-related factors.

Key words: mobile applications, needs assessment, randomized controlled trial, dysmenorrhea, premenstrual syndrome

INTRODUCTION

In 2018, the Android Market and Apple's App Store had 2.1 million and 2 million mobile apps, respectively; by 2020, this will reach 6.1 billion users.^{1,2} The latest report shows that 318 000 of the 6 million apps available for download from the App Store are mobile health (mHealth) apps.³ However, even popular mHealth apps with positive user reviews and high ratings are often inaccurate, and

attempts at regulating medical apps have largely failed.⁴ Thus, healthcare providers must be prepared to answer patient questions about health management using mHealth apps and whether to use these apps.⁵ Moreover, by approaching mHealth app integration using an intentional method, healthcare providers can explore other opportunities to use technology in education, practice, and research.⁶ Accordingly, we developed an app selection method based on user needs

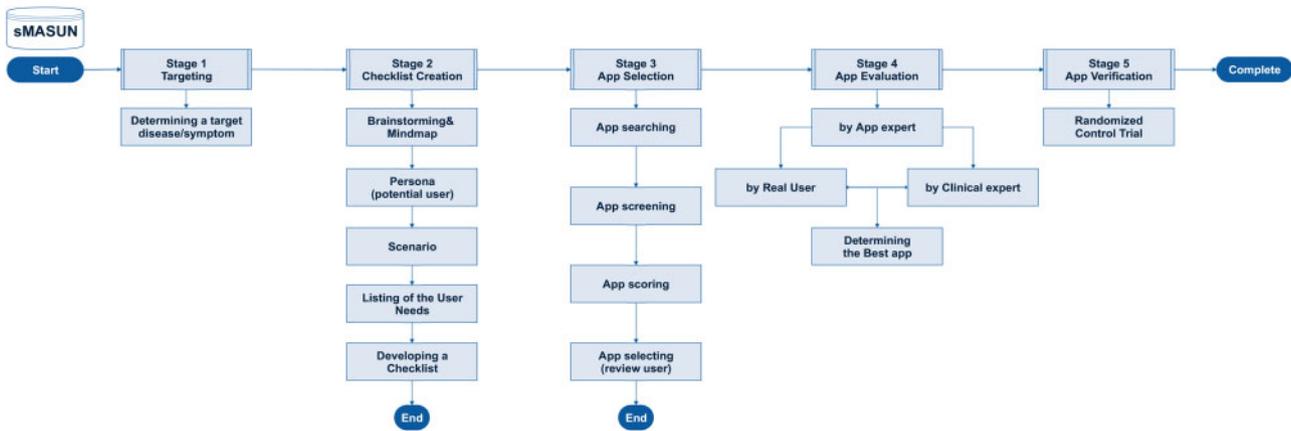


Figure 1. Method of App Selection based on Users' Needs.

(Method of App Selection based on User's Needs [MASUN]) (Figure 1; see the [Supplementary Appendix](#) for the MASUN guidelines).⁷

For this study, we selected participants suffering from dysmenorrhea and premenstrual syndrome (PMS) because 45%-97% of women of reproductive age suffer from dysmenorrhea, which can significantly lower quality of life and sleep quality during the menstruation period.^{8,9} PMS is also common among young women and shares the common risk factors with dysmenorrhea.^{10,11} As many available menstrual apps have not been reviewed by experts, these apps should be evaluated for their usefulness and quality.¹²

The significance of this study is that it will help verify our developed method of app selection. This will, in turn, be valuable for healthcare providers and researchers, providing them with not only a concrete method of choosing the best app for each patient or consumer based on their needs, but also an understanding of how these apps influence changes in consumers' health-related factors.

OBJECTIVES

The purpose of this study is to verify the following research question via a randomized control trial of women with dysmenorrhea and premenstrual syndrome: can menstrual health mobile apps selected based on users' needs lead to changes in health-related factors?

MATERIALS AND METHODS

Study design

This study was a double-blind, pretest-posttest parallel design. In this study, the researcher was not blinded during the intervention; however, random allocation was conducted in a 1:1 manner (intervention group and control group) by a medical informatics expert not involved in this study. Furthermore, cluster and stratified randomization were considered a realistic choice to maximize the number of participants while minimizing bias.¹³ In forming the clusters, acquaintanceship was considered: when several persons already familiar with each other wanted to participate, they were considered as one cluster and randomly assigned to the experimental or control group together. Moreover, menstrual app experience and used app were also considered: when several persons had similar levels of experience in using the same menstrual app, they were divided into 2 groups and randomly assigned.

Participants

The study design, similar literature, and realistic possibilities should be considered when estimating sample size.¹⁴ Accordingly, a sample of 72 individuals was thought best (considering dropout rate). The inclusion criteria were women in their 20s and 30s who used an iPhone (iOS) and reported having dysmenorrhea or PMS. We did not account for dysmenorrhea or PMS severity in recruiting participants because the pain and symptoms for both conditions are highly subjective. Exclusion criteria were taking or planning to take an oral contraceptive during the intervention, having childbirth experience, and having a past diagnosis of pelvic inflammatory disease, endometriosis, or severe emotional or psychiatric illness.

Intervention design

Our intervention was designed using Bandura's social cognitive theory. Bandura presented a theory of the interaction between personal, behavior, and environmental determinants.^{15,16} This theory has been used as a framework for various methodologies in the areas of health promotion and reinforcement because it integrates different factors that determine behavior and can explain behavior change.^{17,18}

The experiment group used the best app chosen through MASUN ([Supplementary Appendix](#)) and the control group used an app with the largest number of downloads and reviews among menstrual app categories for 16 weeks.¹⁹ In the first month, we provided app users with online educational materials and telephone explanations (Figure 2). During these phone explanations (which they did using earphones at our request, as they would need to view the educational materials on their smartphones), the researcher introduced the app to participants and explored the app menus in detail for 10-15 minutes.

To check their app usage, we asked users to send screenshots of specific menus each month. Both groups of users had to send screenshots of a calendar containing participants' records of menstrual or PMS symptoms. The experimental group also needed to send pictures of the following menus from App A: logins, backups, personal alarms, sharing user age information, and evidence-based information about women's health. The control group needed to send pictures of App B's logins, backups, personal alarms, UI selections, and women's health online forums. As a reward for participating in the intervention, we donated a pair of sanitary napkins to low-income girls in participant's name.

Questionnaires

The questionnaire was created online to help researchers save money and time.^{20,21} The questionnaire consisted of the following sections.

App install educational materials

Experimental group

2. Personalized Clue for you



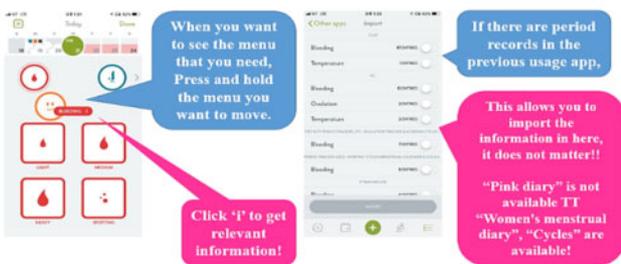
1. Installation, Sign up, and Login



3. To know your body :)



4. Tip for using Clue ~



5. Tip for using Clue ~



So what function can help me?



Every month's progression :)



Figure 2. Educational materials in the first month in the experimental and control groups.

- General characteristics (5 items): Age, occupation, type of smartphone, whether they were experiencing dysmenorrhea or PMS or not
- Dysmenorrhea-related items (11 items): Dysmenorrhea was measured using the visual analog scales used by Kim.²² Two items measured pain in the first and second days of the menstrual period. The remaining 9 items corresponded to various menstrual problems and dysmenorrhea relief methods, which had been extracted, corrected, and supplemented from previous studies.²³
- PMS-related items (10 items): We used the shortened Premenstrual Assessment Form to measure PMS symptoms,²⁴ which is based on the longer Premenstrual Assessment Form.²⁵ The tool comprises 3 parts: emotion (4 items), water congestion (4 items), and pain (2 items). The changes that participants experience a

- week before menstruation is rated on a 6-point Likert-type scale with responses ranging from 1 ("not at all") to 6 ("very severe change"). The higher the score is, the more severe the symptoms of PMS are.
- Social-cognitive factors (11 items):
 - App self-efficacy: Self-efficacy is an important personal factor determining behavior, referring to one's belief that one can accomplish something successfully.¹⁶ We defined app self-efficacy as the belief that one can use an app to successfully manage health problems, including dysmenorrhea and PMS. Four items developed by Jang²⁶ were used. All items use a 7-point Likert-type scale with responses ranging from 1 ("strongly disagree") to 7 ("strongly agree"), with higher

scores indicating higher app self-efficacy. The reliability Jang²⁶ was good (Cronbach's $\alpha=0.82$).²⁶

- App outcome expectancy: Outcome expectancy relates to performance expectations, and is related to actual task performance.²⁷ It also refers to the expected benefits of health habits.²⁸ We defined "app outcome expectancy" as the expectation that using mHealth apps (eg, menstrual apps) will produce positive results. We partially modified Lim's²⁹ 3-item tool, and items are rated on a 5-point Likert-type scale with responses ranging from 1 ("strongly disagree") to 5 ("strongly agree"), with higher scores indicating higher app outcome expectancy.²⁹ The reliability of the tool (Cronbach's α) in Lim's²⁹ study was 0.99.
- App social influence: Social influence predicts individuals' acceptance of information technology.³⁰ We defined app social influence as the individuals' perception that significant others believe that the individual should use health apps. The 4 items developed by Shim³¹ were partially modified to meet the study objectives. This tool utilizes a 5-point Likert-type scale with responses ranging from 1 ("strongly disagree") to 5 ("strongly agree"), with higher scores indicating greater app social influence. In Shim's³¹ study, the Cronbach's α was 0.86.
- Menstrual app experience (7 items): In the prequestionnaire, participants reported whether they used menstrual apps or not; if users had such experience, they completed 5 items on use frequency, menus used, and app influence. In the postquestionnaire, all users completed 2 items about used menus and influence.
- App quality assessment (32 items): All apps were evaluated using the User Version of the Mobile Application Rating Scale, which was validated in previous studies.^{32,33} The app quality rating comprises 4 sections: engagement, functionality, aesthetics, and information. There are also 4 items on app subjective quality (intent to recommend, stimulates repeated use, intention to pay, and overall satisfaction rating), and 6 items assessing the perceived impact of the app in changing the target health behavior. The target health behaviors in this study were menstruation and PMS management behaviors (thus, a total of 12 items were used). The original author authorized use of this tool.

Statistical analyses

The intervention effects were analyzed using *t* test, chi-square tests, analysis of variance (ANOVA), and correlation. If the sample size is large enough (>30), the sample distribution tends to be normal regardless of the data type and violations of the normality assumption do not have severe consequences.^{34,35} Therefore, our analyses were conducted assuming data normality. All analyses used SPSS Statistics 24.0. Two-sided *P* values of <.05 were statistically significant.

Ethical consideration

This study was preapproved by Seoul National University Institutional Review Board (1702/001-011). We collected only minimal personal information for this study and kept all collected data securely in a separate location. When the study is published, all data will be destroyed.

RESULTS

Dysmenorrhea and PMS-related characteristics and menstrual app usage at prequestionnaire

Table 1 shows the characteristics of the 72 users that completed the prequestionnaire. The mean dysmenorrhea pain score of users who

reported having dysmenorrhea was significantly higher (difference = 2.37 points) than was that among users without dysmenorrhea ($t = 2.322$, $P = .023$, $df = 70$; dysmenorrhea group: 5.20 ± 1.68 , $n = 69$; nondysmenorrhea group: 2.83 ± 2.84 , $n = 3$). The mean score for dysmenorrhea relief method was 2.71. Participants' mean PMS score was 29.57 ± 9.88 . Those who reported PMS experience had significantly higher PMS scores ($F = 6.818$, $P = .002$, $df = 71$) than did participants who did not report PMS. PMS experience was significantly related to experience of using apps to manage symptoms ($\chi^2 = 8.147$, $P = .132$).

Fifty-three (73.6%) participants had experience in using menstrual apps. Among them, 8 (15.1%) used apps to check the irregularity of menstruation or PMS. As for how menstrual apps influenced them, 50 (90.3%) users reported that their app only helped them understand their menstrual cycle, while 13 (24.5%) said it helped them understand PMS patterns. Only 1 (1.9%) respondent reported learning about dysmenorrhea or PMS management (Table 1).

Among the 72 app users who answered the prequestionnaire, 63 app users completed the 16-week intervention. Two app users in the experimental group were excluded from the analysis because they did not respond to the survey more than once. Thus, 30 (83.3%) experimental group users and 31 (86.1%) control users were included in the analysis (Figure 3).

Changes in personal outcomes

An ANOVA was conducted to examine the influence of the 2 independent variables (app and time) on app self-efficacy. There were nonsignificant main effects of app ($F = 1.028$, $P > .05$) and time, and a nonsignificant interaction. However, control group's app self-efficacy score decreased by 0.64 points after 16 weeks of using App B, whereas the experimental group's score did not change at all after using App A. For the app quality rating, at baseline, we had participants assess the app quality of menstrual apps they had used before participating in the intervention. The results yielded a significant main effect of app ($F = 16.323$, $P < .01$). That is, App A users showed a significantly higher app quality rating than did App B users. For the experimental group, the average app quality rating increased by 5.6 points over the intervention, while the control group's rating increased by only 3.14 points. A paired *t* test was used to compare app quality rating between the previously used apps and participants' assigned apps. A significant difference was observed in the scores between previously used apps ($n = 21$, mean = 58.00 ± 6.66) and App A ($n = 21$, mean = 65.05 ± 5.88) in the experimental group ($t = -3.829$, $P < .01$). In the control group, there was also a significant difference in scores between previously used apps ($n = 22$, mean = 54.27 ± 7.45) and App B ($n = 22$, mean = 61.32 ± 11.45 ; $t = -2.876$, $P < .01$). There was a significant main effect of app for overall satisfaction ($F = 8.999$, $P < .01$) (Table 2).

Changes in behavioral outcomes

We used an ANOVA to examine the influence of the 2 independent variables (app and time) on app outcome expectancy. The main effect of app was significant ($F = 4.749$, $P < .05$). Regarding the number of menus used, there was only a significant main effect of time ($F = 12.588$, $P < .01$). Post hoc tests revealed that there were significant differences in the number of menus used between previously used apps and the assigned apps (experimental group: $t = -3.0747$, $P < .01$; control group: $t = -7.081$, $P < .01$). Furthermore, in the ex-

Table 1. Descriptive statistics for app users' characteristics and dysmenorrhea and PMS-related variables (N = 72)

Age, y	26.96 ± 4.303
Occupation	
Undergraduate student	17
Graduate student	15
White-collar worker	15
Professional worker	25
Experience of dysmenorrhea	
Yes	69
No	3
Unsure	0
Dysmenorrhea pain score	
First day of period	5.65 ± 2.050
Second day of period	4.54 ± 2.276
Experience of PMS	
Yes	59
No	1
Unsure	13
Mean PMS score	
Yes	31.34
No	11.00
Unsure	21.54
Dysmenorrhea relief method	
Enduring	21 (29.2)
Resting	36 (50.0)
Applying hot pack on abdomen	28 (38.9)
Taking an analgesic	61 (84.7)
Abdomen massage	4 (5.6)
Visiting women's health clinic	1 (1.4)
Menstrual app usage experience	
Have experience	53 (73.6)
Used menus	
Menstruation tracking	52 (98.1)
Ovulation cycle checking	22 (41.5)
Symptom recording	8 (15.1)
Alarm for symptoms	5 (9.4)
Collecting information	0 (0)
Other (eg, intercourse record)	1 (1.9)
Influence	
Understanding the menstruation cycle	50 (90.3)
Getting to know methods for dysmenorrhea relief	1 (1.9)
Understanding the PMS pattern	13 (24.5)
Getting to know method of PMS relief	1 (1.9)
Other (eg, sexual intercourse record)	2 (3.8)
Menstrual app usage experience for managing user's dysmenorrhea and/or PMS	
Have experience	8 (11.1)
No experience	64 (88.9)

Values are mean ± SD or n (%), unless otherwise indicated.
PMS: premenstrual syndrome.

perimental group, the number of menus used increased significantly by the 16th week ($n = 30$, mean = 3.27 ± 1.34) compared with the fourth week ($n = 30$, mean = 2.53 ± 1.28 ; $t = -3.063$, $P < .01$). However, there was no significant increase in the control group ($t = 0.122$, $P > .05$). The most frequently used menu was the that used to check one's menstrual cycle. Regardless of group, all users reported using their apps across all 16 weeks to check their menstrual cycle ($n = 61$, 100%). There were statistically significant differences in the number of recording days per month between the groups ($F = 16.686$, $P < .01$), with a significant increase in the number of recording days between the fourth week and the 16th week in

the experimental group. However, there was a decrease in the control group (Table 2).

Changes in environmental outcomes

An ANOVA was conducted to examine how app and time influenced app social influences. There was a significant main effect of app ($F = 4.242$, $P < .05$). Post hoc tests revealed a significant increase in app social influence between the fourth week (mean = 11.30 ± 3.40) and the 16th week (mean = 12.90 ± 3.34) only in the experimental group ($t = -2.398$, $P < .05$). Additionally, an ANOVA was conducted to examine how the app and time factors influenced intent to recommend. There was a significant main effect of app ($F = 10.974$, $P < .01$) (Table 2).

Changes in dysmenorrhea and PMS-related outcomes

An ANOVA was conducted to examine how app and time influenced dysmenorrhea scores. No significant main effects for either app or time, or a significant interaction effect, were found. The pain reduction was greater for dysmenorrhea score on the second day, decreasing by 1.27 points over the 16-week period in the experimental group and 0.55 points in the control group. Significant differences in the possibility of behavioral/cognitive changes in dysmenorrhea management were also found, with a main effect of app ($F = 5.330$, $P < .05$): in the experimental group, there was a significant increase between the fourth week and the 16th week after using App A ($t = -2.772$, $P < .05$). However, there was no significant increase in the control group ($t = -1.450$, $P > .05$). Moreover, a repeated-measures ANOVA was performed using scores from the fourth, eighth, 12th, and 16th weeks to examine the time effect in more detail. We observed a significant main effect of time ($F = 3.294$, $P < .05$) in the experimental group, but not in the control group ($F = 3.047$, $P > .05$). The number of dysmenorrhea relief methods used increased in both groups; there was no significant difference between the groups ($F = 2.352$, $P > .05$). In the experimental group, the number significantly increased between the fourth week and the 16th week after using App A ($t = -2.191$, $P < .05$). However, there was no significant increase in the control group ($t = 0.441$, $P > .05$) (Table 2).

The principal component analysis of the 14 variables in this study extracted 3 significant components, which together explained 54.87% of the overall variance. Component 1 was related to the menstrual app itself and participants' satisfaction, and was primarily contributed to by app quality rating, overall satisfaction, intent to recommend, and used menu. Component 2 was related to social-cognitive and behavioral factors, and was mainly contributed to by app self-efficacy, app social influence, app outcome expectancy, possibility of behavioral and cognitive changes in dysmenorrhea and PMS management, PMS score, and the number of days with records. Finally, component 3 was related to dysmenorrhea pain and relief method, and was mainly contributed to by the first and second day of menstruation and the number of dysmenorrhea relief methods (Table 3; Figure 4).

DISCUSSION

Dysmenorrhea and PMS-related characteristics

In several studies, primary dysmenorrhea was defined as having a pain score of 3-4 on a 10-item visual analog scale.³⁶ The app users in this study had mean dysmenorrhea pain scores of 5.65 (first day of period) and 4.54 (second day of period) of 10, indicating that many exceeded this diagnostic cutoff. In the present study, the shortened Premenstrual Assessment Form was used to measure PMS and

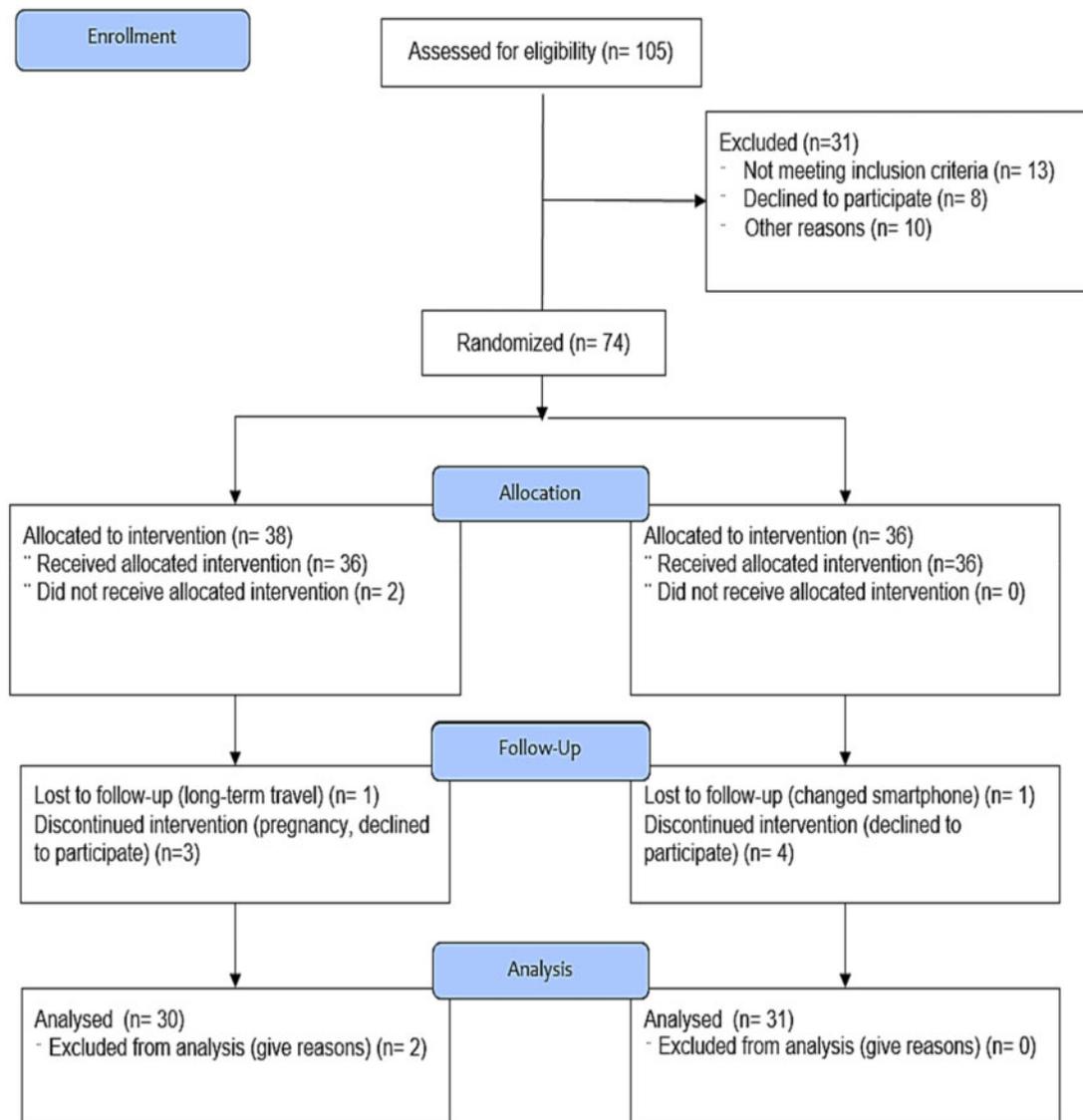


Figure 3. Flow diagram: recruitment and eligibility screening, randomization, follow-up, and analysis.

users' mean PMS score was 29.57. Premenstrual dysphoric disorder can be diagnosed according to a cutoff score of 27 on this scale.²⁴ Users who reported having PMS had a higher PMS score (mean = 31.34). Overall, our results indicate that app users had high scores for both dysmenorrhea and PMS, possibly because these conditions not only influence each other, but also have common influencing factors like poor quality of life.^{10,37,38}

Interestingly, we found no statistically significant differences in dysmenorrhea pain scores; however, experimental group users' PMS score increased after 16 weeks. This might be because these participants obtained greater awareness and understanding of PMS, leading to greater PMS scores. This interpretation is supported by a past study indicating a significant correlation between knowledge of PMS and perception of its severity.³⁹

Findings about menstrual app usage

In the pre-experimental questionnaire, 90.3% of users reported that menstrual apps were helpful for checking period cycles, while

24.5% of users found it helpful for identifying PMS patterns. Only 1.9% of users said that apps helped them learn about dysmenorrhea or PMS management. These results coincide with those of a previous study showing that menstrual app users simply used menstrual apps to track their menstrual cycle.⁴⁰ After 16 weeks, all app users (100%) used the menstrual cycle tracking menu. Compared with a previous study, this proportion was quite high⁴⁰ and suggests that users were more interested in their menstrual cycle because they were already suffering from rather high dysmenorrhea. Additionally, the number of menus used significantly increased over time in both groups. While users did utilize the menstrual apps for recording their menstrual cycles and found it helpful, these apps might be abandoned if predictions about menstrual cycles are inaccurate.⁴¹ Altogether, the menstrual cycle tracking menu seems essential for users' continuing use of the app, as it not only makes users use app but also increases their opportunities to use other menus.

Both groups reported significantly higher app quality ratings for the study apps than for apps used before the study. Thus, both the MASUN-selected best app and the most popular app worldwide

Table 2. Results for personal, behavioral, environmental outcomes, dysmenorrhea, and PMS-related variables

	Experimental Group (n = 30)					Control Group (n = 31)					App Effect	Time Effect	Interaction Effect (App × Time)
	Baseline	4 wk	8 wk	12 wk	16 wk	Baseline	4 wk	8 wk	12 wk	16 wk			
App self-efficacy	20.13	19.57	18.63	20.20	20.13	20.16	19.10	18.52	18.97	19.52	1.028 (.311)	0.738 (.567)	0.136 (.969)
App quality rating	58.38 (n = 21)	64.67	—	—	65.70	54.27 (n = 22)	60.78	—	—	58.84	16.232 ($<.01^a$)	10.070 ($<.01^a$)	0.663 (.517)
Overall satisfaction	3.43 (n = 21)	3.97	3.93	3.87	3.93	3.41 (n = 22)	3.58	3.68	3.42	3.68	8.999 (.003 ^a)	2.263 (.063)	0.575 (.681)
App outcome expectancy	11.17	10.93	11.03	11.23	11.80	10.74	10.61	10.35	10.74	10.94	4.749 (.030 ^a)	0.838 (.502)	0.144 (.966)
Used menu	1.86 (n = 21)	2.53	2.93	3.17	3.33	1.36 (n = 22)	2.74	2.94	3.10	2.90	1.145 (.286)	12.588 ($<.01^*$)	0.807 (.522)
Number of days with records	—	5.67	7.13	7.27	7.10	—	3.58	3.81	3.87	3.35	16.686 ($<.01^a$)	0.305 (.822)	0.223 (.881)
App social influence	11.43	11.30	12.03	12.83	12.90	10.55	11.06	11.48	12.06	11.29	4.242 (.040 ^a)	1.930 (.105)	0.338 (.825)
Intent to recommend	2.81	3.47	3.27	3.47	3.50	2.86	3.06	2.94	2.74	3.03	10.974 ($<.01^a$)	1.682 (.154)	1.104 (.355)
First day of menstruation	5.53	5.50	5.33	5.57	5.77	5.74	5.51	5.83	5.71	5.45	0.233 (.629)	0.043 (.996)	0.334 (.855)
Second day of menstruation	4.90	5.20	4.60	4.63	4.50	4.55	4.68	4.45	5.06	4.90	0.022 (.883)	0.300 (.878)	0.578 (.679)
Possibility of behavioral and cognitive changes in Dysmenorrhea management	15.67 (n = 21)	20.60	21.57	22.10	22.43	15.41 (n = 22)	20.23	20.06	19.61	21.19	5.330 (.022 ^a)	15.782 ($<.01^a$)	0.642 (.633)
Number of dysmenorrhea relief methods	2.20	2.27	2.47	2.53	2.80	2.06	2.41	2.25	2.13	2.45	2.352 (.126)	1.631 (.166)	0.632 (.640)
PMS score	30.40	32.27	31.83	32.93	32.40	30.39	30.16	29.84	30.13	30.71	2.405 (.122)	0.156 (.960)	0.171 (.953)
Possibility of behavioral and cognitive changes in PMS management	15.90 (n = 21)	21.73	22.60	22.97	23.10	13.72 (n = 22)	20.83	19.97	19.94	21.13	14.680 ($<.01^a$)	19.090 ($<.01^a$)	0.454 (.769)

P values are presented in parentheses.

PMS: premenstrual syndrome.

^a $P < .05$.

were better than others.¹⁹ However, the groups showed significant differences in app quality ratings: App A had a higher rating than did App B. Interestingly, over the 16-week period, App A's rating increased but App B's rating decreased. Additionally, App A had significantly higher overall satisfaction scores than did App B. The means of the number of menus used and recording days both increased significantly by the 16th week, but only in the experimental group. Specifically, they recorded their symptoms around 3-4 days more than did control group users. It is exceedingly important that apps provide information on users' symptoms or promote treatment adherence via reminders and notifications.⁴² Moreover, previous studies have shown that app social influence is positively related to continued app usage intention, and has a direct positive impact on perceived app usefulness.⁴³ Women's continuance decisions were especially impacted by social influences.⁴⁴ App social influence and intention to recommend to other persons significantly differed by app, with the experimental group having higher mean scores for both variables than the control group. In other words, our findings suggest that if app users are satisfied with an app selected with consideration of their needs, they not only rate the app as higher in quality but also use the app for longer and spread it to other users.

Interestingly, there was a large gap in the number of recording days within the same group: Some users recorded on fewer than 3

days/month (9 users in the experimental group, 12 users in the control group), while others recorded approximately every day (more than 20 days; 3 users in the experimental group, 1 user in the control group). These results closely agree with those of a study examining apps for reporting users' pain.⁴⁵ In the future, researchers should examine the factors that prevent users from recording.

Changes in personal, behavioral, and environmental outcomes

Electronic health tools such as mHealth apps for tracking users' symptoms can have a positive effect on self-efficacy.⁴⁶ We found that although the number of recording days increased significantly in the experimental group, app self-efficacy did not. This is possibly because it is difficult to elevate self-efficacy significantly using mobile technology-related healthcare interventions.⁴⁷ However, features that stimulate self-efficacy can still help promote continuous mHealth app usage and enact changes in health behavior.⁴⁸⁻⁵¹ These results coincide well with those found in our study. There was a slight increase in app self-efficacy score, and significant correlations were observed between component 3 and app self-efficacy, possibility of behavioral and cognitive changes in dysmenorrhea and PMS management, and the number of days with records. Particularly, in

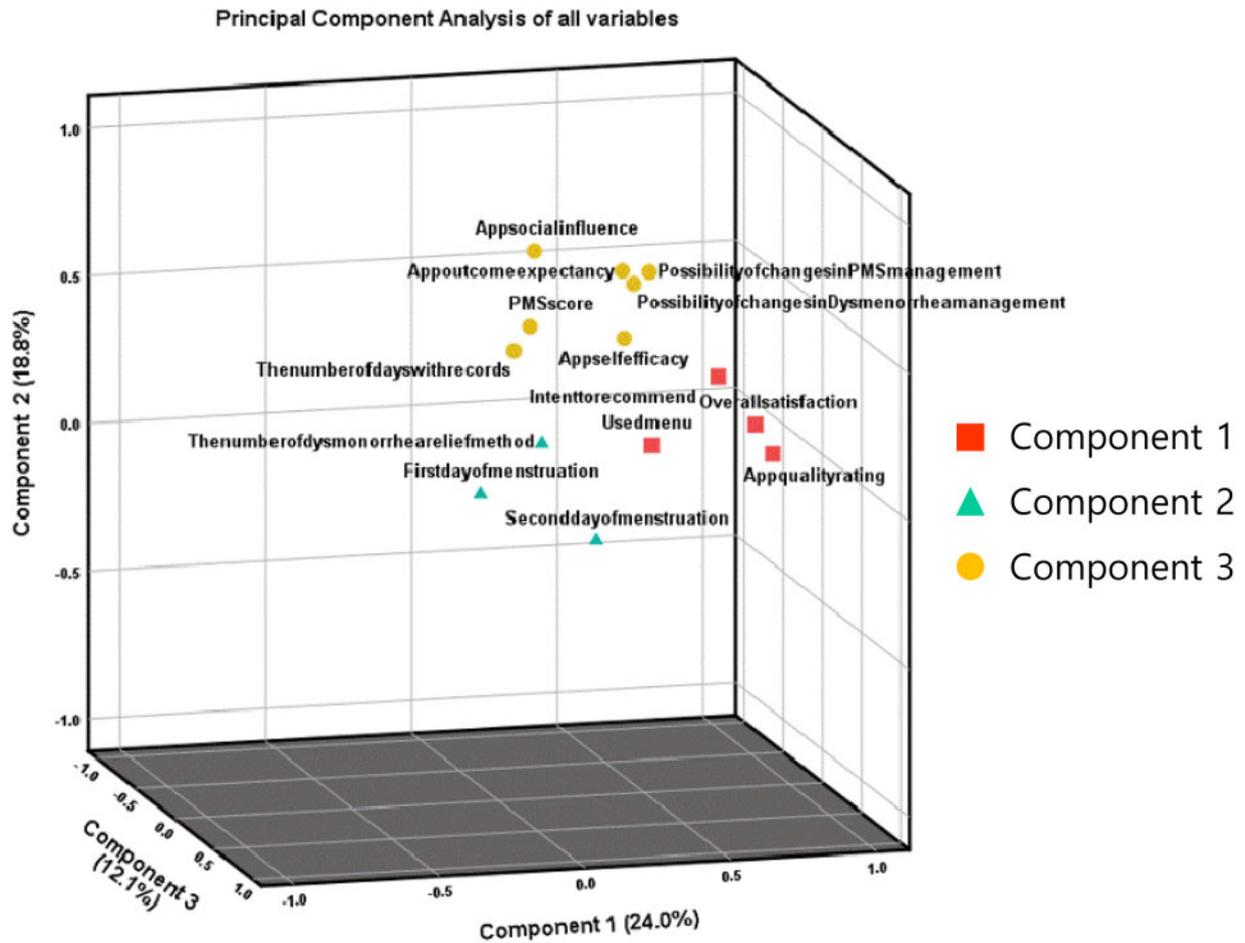


Figure 4. Variables contained in the 3 principal component dimensions with the variance in the rotation sums of squared loadings.

Table 3. Principal component analysis of all variables

Variable	Factor Loading		
	Component 1	Component 2	Component 3
App quality rating	.907 ^a		
Overall satisfaction	.884 ^a		
Intent to recommend	.756 ^a		
Used menu	.416 ^a		.406
App social influence		.770 ^a	
App outcome expectancy	.421	.662 ^a	
Possibility of behavioral and cognitive changes in PMS management	.584	.595 ^a	
PMS score		.580 ^a	.419
Possibility of behavioral and cognitive changes in dysmenorrhea management	.534	.556 ^a	
App self-efficacy	.347	.500 ^a	.311
The number of days with records		.420 ^a	
Second day of menstruation			.650 ^a
First day of menstruation			.621 ^a
The number of dysmenorrhea relief methods			.528 ^a
Eigenvalue	3.36	2.63	1.69
Variance, %	24.02	18.80	12.05
Cumulative variance, %	24.02	42.82	54.87

^aVariable contained in the component.

PMS: premenstrual syndrome.

the experimental group, significant changes after 16 weeks were found in the possibility of both types of change (behavioral and cognitive changes in dysmenorrhea and PMS management). The potential reason is that apps with features that target both internal factors (eg, self-efficacy, app outcome expectancy) and external factors (eg, disease information, social networking, and user compatibility) tend to encourage greater behavioral change.⁵¹ Unlike App B (used by the control group), App A (used by the experimental group) provided information on dysmenorrhea and PMS, as well as had a social networking function enabling users to exchange information on their menstrual cycle with other users.

In a previous study, outcome expectancy did not positively influence young users' behavioral intention to use educational computing technology.⁵² However, significant increases in app outcome expectancy were also observed in the experimental group and significant correlations were observed between app outcome expectancy and the possibility of behavioral and cognitive change in dysmenorrhea and PMS management. This suggests that as users expect an mHealth app to have positive outcomes, they tend to be more motivated to perform the related health behavior.⁴⁸

Furthermore, mHealth apps ideally help users manage symptoms and improve the health outcomes of people with chronic symptoms.⁵³ Indeed, component 3 in the principal component analysis grouped pain of dysmenorrhea and pain relief method variables into a single component. Interestingly, the number of dysmenorrhea relief methods in the experimental group increased significantly. This suggests that mHealth apps with effective menus for managing symptoms that match users' needs could help these users engage in more health relief methods.

Strengths of user needs consideration

Given the increased interest in mHealth apps, it is necessary to optimize the process by healthcare providers choose and utilize such apps for referrals.⁵⁴ Boudreaux et al's⁵⁴ study revealed 7 strategies for choosing mHealth apps: (1) review of scientific literature; (2) search app clearinghouse websites; (3) search app stores; (4) check app descriptions, ratings, and reviews; (5) consult the clinical experts and patient networks on social media; (6) pilot the apps; and (7) get feedback from patients. The MASUN is consistent with steps 3, 4, 6, and 7. Furthermore, in both our and Boudreaux et al's⁵⁴ studies, researchers and users were recruited using social network services. By integrating mHealth apps into clinical practices and interventions, healthcare providers and researchers can successfully improve client and consumer outcomes, particularly satisfaction.⁶ This is why, when selecting the best app for target users, it is necessary to consider not only user's needs, but also app and clinical experts' opinions. MASUN provides a useful way of integrating expert opinions with the needs of various users through stages 2-4 (Figure 1 and Supplementary Appendix; also, a simpler version of MASUN is available upon request from JL).

CONCLUSION

This study is significant because it helped verify a new method of app selection based on users' needs via a randomized controlled trial. When using the app that best reflected their needs, users recorded their symptoms more often and reported higher app social influence, intent to recommend to others, app quality rating, and overall satisfaction. The results not only benefit the selection of menstrual apps,

but also confirm how mHealth apps can benefit health behavior (as evidenced by the fact that the app selected using MASUN promoted users' reported possibility of behavioral and cognitive changes in dysmenorrhea and PMS management). Taken together, our findings indicate that healthcare providers or researchers should fully consider the target app users' needs.

Based on the results of this study, 2 follow-up studies are underway. In the first, we aim to simplify and standardize a new method to more easily select the best app for users with other health problems. We are reviewing the MASUN using the Delphi technique with a multidisciplinary team of experts. We also intend to adapt the MASUN into an online format so that this new method can be utilized internationally. In the second study, we are designing a menstrual app user interface that reflects the needs of particular users. This could be used by developers or researchers who want to develop apps that enable effective menstrual management, including dysmenorrhea and PMS.

FUNDING

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education (NRF-2015R1D1A1A01061329 [JK], NRF-2018R1C1B5030802 [JL]).

AUTHOR CONTRIBUTIONS

Both authors had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Both authors were involved in study concept and design; JL was involved in acquisition of data; JL was involved in statistical analysis of data; both authors were involved in interpretation of data; both authors drafted the manuscript; both authors were involved in critical revision of the manuscript for important intellectual content; Both authors obtained funding; JK provided administrative, technical, and material support; JL supervised the study.

ACKNOWLEDGEMENTS

The authors are grateful to Dr Chan Hee Park, who allocated the randomization group and all app users who participated in the intervention, and to Dr Soohyung Kim, who gave advice on statistical analysis.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. Statista. Number of apps available in leading app stores as of 3rd quarter 2018. <https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/>.
2. Galinina O, Turlikov A, Andreev S, Koucheryavy Y. Improving reliability of replicated message delivery in cellular machine-type communications. In: *2016 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*; 2016: 106–10.
3. Aitken M, Clancy B, Nass D. *The Growing Value of Digital Health: Evidence and Impact on Human Health and the Healthcare System*. Durham, NC: IQVIA Institute for Human Data Science; 2017.
4. Plante TB, O'Kelly AC, Macfarlane ZT, et al. Trends in user ratings and reviews of a popular yet inaccurate blood pressure-measuring smartphone app. *J Am Med Inform Assoc* 2018; 25 (8): 1074–9.

5. Torous JB, Chan SR, Gipson SYT, *et al.* A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. *Psychiatr Serv* 2018; 69 (5): 498–500.
6. Arbour MW, Stec MA. Mobile applications for women's health and midwifery care: a pocket reference for the 21st century. *J Midwifery Womens Health* 2018; 63 (3): 330–4.
7. Lee J, Kim J. Method of app selection for healthcare providers based on consumer needs. *Comput Inform Nurs* 2018; 36 (1): 45–54.
8. Latthe P, Latthe M, Say L, *et al.* WHO systematic review of prevalence of chronic pelvic pain: a neglected reproductive health morbidity. *BMC Pub Health* 2006; 6 (1): 177.
9. Iacovides S, Avidon I, Baker FC. What we know about primary dysmenorrhea today: a critical review. *Hum Reprod Update* 2015; 212 (6): 762–78.
10. Ju H, Jones M, Mishra GD. Premenstrual syndrome and dysmenorrhea: symptom trajectories over 13 years in young adults. *Maturitas* 2014; 78 (2): 99–105.
11. Ju H. *The Magnitude, Long-Term Trend and Lifestyle Risk Factors of Dysmenorrhea and Premenstrual Syndrome* [PhD thesis]. Herston, Australia, School of Public Health, the University of Queensland; 2016. <https://doi.org/10.14264/uql.2016.349>
12. Moglia ML, Castano PM. A review of smartphone applications designed for tracking women's reproductive health [111]. *Obstetr Gynecol* 2015; 125: 41S.
13. Glynn RJ, Brookhart MA, Stedman M, *et al.* Design of cluster-randomized trials of quality improvement interventions aimed at medical care providers. *Med Care* 2007; 45 (Suppl 2): S38–43.
14. Pye V, Taylor N, Clay-Williams R, *et al.* When is enough, enough? Understanding and solving your sample size problems in health services research. *BMC Res Notes* 2016; 9 (1): 1.
15. Bandura A. Social cognitive theory of mass communication. *Media Psychol* 2001; 3 (3): 265–99.
16. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* 1977; 84 (2): 191–215.
17. Clark NM, Zimmerman BJ. A social cognitive view of self-regulated learning about health. *Health Educ Behav* 2014; 41 (5): 485–91.
18. Stacey FG, James EL, Chapman K, *et al.* A systematic review and meta-analysis of social cognitive theory-based physical activity and/or nutrition behavior change interventions for cancer survivors. *J Cancer Surviv* 2015; 9 (2): 305–38.
19. Aitken M, Lyle J. *Patient Adoption of mHealth: Use, Evidence and Remaining Barriers to Mainstream Acceptance*. Parsippany, NJ: IMS Institute for Healthcare Informatics; 2015.
20. Wright KB. Researching Internet-based populations: advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *J Comp Mediated Commun* 2005; 10 (3): JCMC1034.
21. Waclawski E. How I use it: survey monkey. *Occup Med (Lond)* 2012; 62 (6): 477.
22. Kim EA. *Effects of Tourmaline Gemstone Therapy on Dysmenorrhea, menstrual Pain and Prostaglandins of the Female University Students*. Seoul, South Korea: Chungang University; 2007.
23. Shin EH. *Effects of Heated Red Bean Pillow and Aromatherapy for 20–30 Years Old Women's Menstrual Pain and Menstrual Discomforts*. Seoul, South Korea: Dongduk Women's University; 2014.
24. Allen S, McBride C, Pirie P. The shortened premenstrual assessment form. *J Reprod Med* 1991; 36 (11): 769–72.
25. Halbreich U, Endicott J, Schacht S, *et al.* The diversity of premenstrual changes as reflected in the Premenstrual Assessment Form. *Acta Psychiatr Scand* 1982; 65 (1): 46–65.
26. Jang S. *Factors Influencing the Use Intention of Smartphone Application - Comparative Study of College Students in Korea and China*. Cheongju, South Korea: Chungbuk National University; 2012.
27. Pintrich PR. A conceptual framework for assessing motivation and self-regulated learning in college students. *Educ Psychol Rev* 2004; 16 (4): 385–407.
28. Bandura A. Health promotion by social cognitive means. *Health Educ Behav* 2004; 31 (2): 143–64.
29. Lim JS. *A Study on the Effect of the Introduction Characteristics of Cloud Computing Services on the Performance Expectancy and the Intention to Use: Focusing on the Innovation Diffusion Theory*. Yongin, South Korea: Dankook University; 2012.
30. Thompson RL, Higgins CA, Howell JM. Personal computing: toward a conceptual model of utilization. *Mis Q* 1991; 15 (1): 125–43.
31. Shim YB. *Factors Related to the Intent to Use the Medical Application (M-APP) Smart Phone of Hospital Employees*. Wonju, South Korea: Yonsei University, Graduate School of Health and Environment; 2011.
32. Stoyanov SR, Hides L, Kavanagh DJ, *et al.* Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015; 3 (1): e27.
33. Zhou L, Bao J, Parmanto B. Systematic review protocol to assess the effectiveness of usability questionnaires in mhealth app studies. *JMIR Res Protoc* 2017; 6 (8): e151.
34. Field A. *Discovering Statistics Using SPSS*. Thousand Oaks, CA: Sage; 2009.
35. Pallant J. *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS*. London: Open University Press; 2011.
36. Daily JW, Zhang X, Kim DS, Park S. Efficacy of ginger for alleviating the symptoms of primary dysmenorrhea: A systematic review and meta-analysis of randomized clinical trials. *Pain Med* 2015; 16 (12): 2243–55.
37. Washington MJ, Brown C, Ling FW. *Premenstrual Syndrome Premenstrual Dysphoric Disorder*. Cambridge, UK: Cambridge University Press; 2015.
38. Lustyk MK, Widman L, Paschane A, *et al.* Stress, quality of life and physical activity in women with varying degrees of premenstrual symptomatology. *Women Health* 2004; 39 (3): 35–44.
39. Dadi Givshad R, Nourani SS, Esmaily H. The relationship of perceived severity of premenstrual syndrome with knowledge, attitude and recorded severity of syndrom by a daily calendar among university students in Iran. *J Midwifery Reprod Health* 2016; 4 (1): 522–9.
40. Bretschneider RA. A goal-and context-driven approach in mobile period tracking applications. In: Antona M, Stephanidis C, eds. *Universal Access in Human-Computer Interaction. Access to Learning, Health and Well-Being, UAHCI 2015*. Cham, Switzerland: Springer; 2015: 279–87.
41. Epstein DA, Lee NB, Kang JH, *et al.* Examining menstrual tracking to inform the design of personal informatics tools. *Proc SIGCHI Conf Hum Factor Comput Syst* 2017; 2017: 6876–88.
42. Nightingale R, Hall A, Gelder C, *et al.* Desirable components for a customized, home-based, digital care-management app for children and young people with long-term, chronic conditions: A qualitative exploration. *J Med Internet Res* 2017; 19 (7): e235.
43. Wang EST, Chou NPY. Examining social influence factors affecting consumer continuous usage intention for mobile social networking applications. *Int J Mobile Commun* 2016; 14 (1): 43–55.
44. Lu J. Are personal innovativeness and social influence critical to continue with mobile commerce? *Int Res* 2014; 24 (2): 134–59.
45. Graul A, Zhang X, Schmitz K, *et al.* The frequency of self-reported symptoms of lymphedema over 5 years in patients with uterine carcinoma. *Gynecol Oncol* 2017; 145: 205–6.
46. Wolf A, Fors A, Ulin K, *et al.* An eHealth diary and symptom-tracking tool combined with person-centered care for improving self-efficacy after a diagnosis of acute coronary syndrome: a substudy of a randomized controlled trial. *J Med Internet Res* 2016; 18 (2): e40.
47. Pfaeffli Dale L, Dobson R, Whittaker R, *et al.* The effectiveness of mobile-health behaviour change interventions for cardiovascular disease self-management: A systematic review. *Eur J Prev Cardiol* 2016; 23 (8): 801–17.
48. Bandura A. *Self-efficacy: The Exercise of Control*. London: Macmillan; 1997.
49. Holroyd KA, Creer TL. *Self-management of Chronic Disease: Handbook of Clinical Interventions and Research*. Orlando, FL: Academic Press; 1986.

50. Park M, Yoo H, Kim J, Lee J. Why do young people use fitness apps? Cognitive characteristics and app quality. *Electron Commer Res* 2018; 18 (4): 755–61.
51. Fitzgerald M, McClelland T. What makes a mobile app successful in supporting health behaviour change?. *Health Educ J* 2017; 76 (3): 373–81.
52. Ratten V. Cloud computing: A social cognitive perspective of ethics, entrepreneurship, technology marketing, computer self-efficacy and outcome expectancy on behavioural intentions. *Austral Market J* 2013; 21 (3): 137–46.
53. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. *J Med Internet Res* 2016; 18 (5): e97.
54. Boudreaux ED, Waring ME, Hayes RB, et al. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. *Behav Med Pract Policy Res* 2014; 4 (4): 363–71.
55. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res Nurs Health* 2007; 30 (4): 459–67.

APPENDIX. MASUN guidelines

Tasks	Descriptions	Stakeholders	Estimated Time
Stage 1. TARGETING			
1. Targeting	Determine the [target disease and/or symptom]	Researcher who participated in the current study	
Stage 2. CHECK LIST CREATION			
2. Brainstorming and Mind Mapping	Conducting group brainstorming sessions on the [target disease and/or symptom] and designing mind map about brainstorming	Individual (healthcare provider or researcher) who had experience/majored in informatics (eg, medical informatics, nursing informatics) (≥2 groups, group ≥2 individuals)	1 hour
3. Persona and Scenario	Design a potential user (persona) who uses an app to manage the [target disease and/or symptom] (≥2 personas) Create a scenario for each persona who uses an app to manage the [target disease and/or symptom] (≥2 scenarios)	*The same groups as the previous stage	3 hours
4. User Needs	List of user's needs (requirements) derived from each scenario of each persona who uses an app to manage the [target disease and/or symptom], Write the list followed by order of priority	*The same groups as the previous stage	1 hour
5. Developing Checklist	Draft version of checklist by table format: Compile all of the user's needs (requirements) included in the scenario into a checklist	*Same individual can do this stage with stage 1 Researcher who participated in the current study	1 hour
	Modifying checklist: Select items (user's needs, requirements) based on a content validity index of >0.75.	† Different individual with all previous stages Individual (Healthcare provider or researcher) who had experience/majored in informatics (eg, medical informatics, nursing informatics) (≥3 individuals) [55]	1 day
	Final version of checklist: Allocate partial score for each item. <i>If the user's needs are "recording the pain in the app," allocate 0 points to the item in the checklist if there is no pain recording menu, 1 point if there is a pain recording menu with a dichotomous (Yes/No) scale, and 2 points if there is a pain recording menu with Likert or score recording system.</i>	*Same individual can do this stage with previous stage (second version of checklist) Individual who majored in medical informatics or nursing informatics (≥3)	1 day
Stage 3. APP SELECTION			
6. App Searching	Search terms: Found synonyms for [target disease and/or symptom]. <i>If a researcher wants to search for menstrual apps, he/she needs to search using terms such as menstruation, period, and menstrual cycle.</i> App search: Searched for apps in 1 or more app stores using the derived search terms.	Same individual can do this stage with stage 1 or 6 Researcher who participated in the current study	2–3 hours
		*Same individual can do this stage with stage 1 or 6 or 8 Researcher (≥2)	1 day for 1 search term

(continued)

***Appendix.** continued

Tasks	Descriptions	Stakeholders	Estimated Time
7. App Screening	Deduplication: Compared the lists of the searched apps and compiled them into a single list. If the app name and developer were the same, they are considered the same app.	*Same individual as the previous stage Researcher (≥ 2)	1-2 days
	Remove all apps unrelated to the [target disease and/or symptom] based on the app name or description.	*Same individual as the previous stage Researcher (≥ 2)	1-2 days
8. App Scoring	Score the apps using the checklist after using app at least 10 mins (Independently).	*Same individual as the previous stage Researcher (≥ 2)	1 day for 30 apps
	Crosscheck: Compare the scores of the top 100 apps given by each researcher. If the total scores of the app or the scores given for a specific item in the checklist differ between the researchers, the researchers should review the app and discuss it to reach a consensus on the score.	*Same individual as the previous stage Researcher (≥ 2)	1-2 days
9. App Selecting	Check a review user's needs (the app rating and the number of reviews) for each app in app store with the top 5 checklist scores. A given checklist score might include multiple apps.	*Same individual as the previous stage Researcher	1 hour
	When multiple apps receive the same score, the app with the highest rating should be selected.	*The same person as the previous stage Researcher	1 hour
	When multiple apps receive the same score and app rating, the app with the highest number of reviews should be selected.	*The same person as the previous stage Researcher	1 hour
Stage 4. APP EVALUATION			
10. App Evaluation	Evaluate the candidate apps using an assessment tool like MARS (Mobile Application Rating Scale) after using each app for at least 10 mins.	†App expert with experience in developing and/or designing a health-related app (≥ 5 individual)	1 day
	Evaluate the candidate apps using an assessment tool like the user-MARS after using each app for at least 10 mins.	†Clinical expert with experience in the treatment of and/or research on [target disease and/or symptom] patients/consumers (≥ 5 individual)	1 day
	Evaluate the candidate app through focus group interviews (real users). The interview questions should be structured around an assessment tool like the user-MARS. An evaluation (preferred app or not) should be conducted after using each app for at least 10 mins.	†Person who has experienced the [target disease and/or symptom] (≥ 3 groups, group ≥ 2 individual)	1 day for 1 group
11. Determining the Best App	Among the preferred apps, the app with the highest score given by the app and clinical experts is selected as the best app.	Same individual can do this stage with stage 1 or 6 or 8 Researcher who participated in the current study	1 day
Stage 5. APP VERIFICATION			
12. Randomized Controlled Trial	Randomized controlled trial is conducted. An intervention group use the best app selected in the previous stage.	† Person who has experienced the [target disease and/or symptom] (≥ 2 groups)	

*Person who participated in previous stage can participate in this step.

†Person who participated in the previous stage cannot participate in this step. In other words, a new participant/researcher should perform this step.