

Original Paper

# Using Clinician-Patient WeChat Group Communication Data to Identify Symptom Burdens in Patients With Uterine Fibroids Under Focused Ultrasound Ablation Surgery Treatment: Qualitative Study

Jiayuan Zhang<sup>1</sup>, MSc; Wei Xu<sup>2</sup>, PhD; Cheng Lei<sup>2</sup>, MSN; Yang Pu<sup>1</sup>, PhD; Yubo Zhang<sup>1</sup>, MSc; Jingyu Zhang<sup>1</sup>, PhD; Hongfan Yu<sup>2</sup>, MSc; Xueyao Su<sup>2</sup>, MSc; Yanyan Huang<sup>2</sup>, MSc; Ruoyan Gong<sup>2</sup>, MSc; Lijun Zhang<sup>2</sup>, MPH; Qiuling Shi<sup>1,2</sup>, MD, PhD

<sup>1</sup>State Key Laboratory of Ultrasound in Medicine and Engineering, College of Biomedical Engineering, Chongqing Medical University, Chongqing, China

<sup>2</sup>School of Public Health, Chongqing Medical University, Chongqing, China

**Corresponding Author:**

Qiuling Shi, MD, PhD  
School of Public Health  
Chongqing Medical University  
No 1, Medical College Road  
Yuzhong District  
Chongqing, 400016  
China  
Phone: 86 18290585397  
Email: [qshi@cqmu.edu.cn](mailto:qshi@cqmu.edu.cn)

## Abstract

**Background:** Unlike research project-based health data collection (questionnaires and interviews), social media platforms allow patients to freely discuss their health status and obtain peer support. Previous literature has pointed out that both public and private social platforms can serve as data sources for analysis.

**Objective:** This study aimed to use natural language processing (NLP) techniques to identify concerns regarding the postoperative quality of life and symptom burdens in patients with uterine fibroids after focused ultrasound ablation surgery.

**Methods:** Screenshots taken from clinician-patient WeChat groups were converted into free texts using image text recognition technology and used as the research object of this study. From 408 patients diagnosed with uterine fibroids in Chongqing Haifu Hospital between 2010 and 2020, we searched for symptom burdens in over 900,000 words of WeChat group chats. We first built a corpus of symptoms by manually coding 30% of the WeChat texts and then used regular expressions in Python to crawl symptom information from the remaining texts based on this corpus. We compared the results with a manual review (gold standard) of the same records. Finally, we analyzed the relationship between the population baseline data and conceptual symptoms; quantitative and qualitative results were examined.

**Results:** A total of 408 patients with uterine fibroids were included in the study; 190,000 words of free text were obtained after data cleaning. The mean age of the patients was 39.94 (SD 6.81) years, and their mean BMI was 22.18 (SD 2.78) kg/m<sup>2</sup>. The median reporting times of the 7 major symptoms were 21, 26, 57, 2, 18, 30, and 49 days. Logistic regression models identified preoperative menstrual duration (odds ratio [OR] 1.14, 95% CI 5.86-6.37;  $P=0.009$ ), age of menopause (OR -1.02, 95% CI 11.96-13.47;  $P=0.03$ ), and the number (OR 2.34, 95% CI 1.45-1.83;  $P=0.04$ ) and size of fibroids (OR 0.12, 95% CI 2.43-3.51;  $P=0.04$ ) as significant risk factors for postoperative symptoms.

**Conclusions:** Unstructured free texts from social media platforms extracted by NLP technology can be used for analysis. By extracting the conceptual information about patients' health-related quality of life, we can adopt personalized treatment for patients at different stages of recovery to improve their quality of life. Python-based text mining of free-text data can accurately extract symptom burden and save considerable time compared to manual review, maximizing the utility of the extant information in population-based electronic health records for comparative effectiveness research.

(JMIR Form Res 2023;7:e43995) doi: [10.2196/43995](https://doi.org/10.2196/43995)

**KEYWORDS**

social media; group chats; text mining; free texts; symptom burdens; WeChat; natural language processing; NLP

## Introduction

Uterine fibroids (UFs) are the most common benign diseases of the reproductive tract in women of reproductive age, and their prevalence has been increasing in recent decades, ranging from 50% to 77% [1]. UFs are commonly observed in women of childbearing age, widowhood, and uncoordinated sexual life [2]. They can cause abdominal masses, increased menstrual flow, dizziness, fatigue, frequent urination, and secondary dysmenorrhea. Therefore, alleviating adverse symptoms and improving the quality of life are the primary purposes of UF treatment.

Focused ultrasound ablation surgery (FUAS) is an emerging technology for the treatment of UFs without any visible incision [3-6]. The median hospitalization time for patients is 2 days [7]. Postoperative complications may occur from hospitalization to several months after surgery, and a substantial proportion of patients may experience symptoms related to disease recurrence. Frequent monitoring using standardized patient-reported outcome (PRO) tools could improve patients' quality of life in both open and minimally invasive surgical practice [8]. Whether this approach works similarly in noninvasive surgery settings is still unknown because of the lack of understanding of what patients experience throughout recovery and disease recurrence.

Unlike those with malignant diseases, patients with UFs are not rushed to seek professional medical advice when a symptom bothers them, and they tend to turn to the internet for help [9]. Information from Facebook or Twitter [1-4] has served as a source for identifying patients' experiences and health care needs during the perioperative phase. In China, the most widely used social media platform is WeChat, covering approximately 1.2 billion internet users of all ages [10]. Clinician-patient WeChat groups are popular in all types of hospitals and are an effective instrument for patient care after surgery. Unlike discussion forums open to all viewers on other social media, clinician-patient WeChat groups are established by medical staff, and since only patients treated by those medical staff are granted access, the groups are relatively private. In these groups, patients may report their symptoms, consult with doctors for their medical examination reports, and seek guidance for their next clinical visits. Medical staff respond within 24 hours, with explanations or suggestions. A patient may share their experience with anything related to their diseases and treatments with other patients because they are grouped for similar reasons. Because of insufficient medical personnel resources and heavy medical workloads, clinician-patient communication in China is extremely brief in clinics, and patients might feel that their doctors do not care about them. Thus, clinician-patient WeChat groups have been widely accepted as a complementary approach to relieve tension between doctors and patients [5].

However, the use of information from clinician-patient WeChat groups has been limited. As a nonprofessional communication platform, in addition to disease-related text, the WeChat group chat is also filled with a large amount of noisy data, which

interferes with the artificial intelligence-based learning of natural language processing (NLP) [11], causing inaccurate data extraction of disease- and treatment-related information. To appropriately use and integrate information from clinician-patient WeChat communication, specific algorithms for this type of natural language are necessary.

Therefore, with the availability of patients' postdischarge experience from the clinician-patient WeChat groups and disease and treatment information from the electronic health records (EHRs), we aimed to develop an algorithm for using NLP techniques to extract post-FUAS symptoms experienced by patients from the clinician-patient WeChat group text, describe symptom burdens before and after FUAS for UFs, and determine the risk factors of postoperative symptom burdens [6-9].

## Methods

### Study Design

This study used a method that consisted of WeChat free-text NLP and the analysis of quantitative data from the EHR.

### Ethics Approval

The Institutional Review Board of Chongqing Haifu Hospital approved the protocol (approval number: 2020-003). Informed consent was waived for the retrospective design. The registration number of the China Clinical Trial Center was ChiCTR2200056735.

### Data Source

We used ABBYY FineReader 14 to extract the texts from the screenshots of clinician-patient WeChat groups, including the communication between patients and attending physicians from January 2010 to December 2020. We obtained screenshots of clinician-patient WeChat groups from the patient management project after FUAS treatment, which consisted of a continuous report from patients on their state and the clinician's response to patients, as shown in Figure S1 in [Multimedia Appendix 1](#).

### Data Cleaning

We filtered the WeChat group chat texts using the following keywords: “肌瘤” (“fibroid” in Chinese), “子宫肌瘤” (“uterine fibroid” in Chinese), “多发子宫肌瘤” (“multiple fibroids” in Chinese), “多肌” (“multiple muscle” in Chinese), and “子宫肌” (“uterine muscle” in Chinese). For text mining, we removed all the emojis and created a script to remove URLs and texts of hospital examination records. To ensure that the data only include the text related to the research topic, we cleaned the texts by removing words and characters that were of little or no analytical value (eg, “早上” [“morning” in Chinese] and “下午” [“afternoon” in Chinese]). Finally, only the communication texts of patients related to UFs were preserved [12-14]. The cleaned data were randomly divided into a training data set (30%) and a validation data set (70%).

## Manual Coding

We convened an expert panel of gynecologists, computer programmers, and medical informatics professionals, which categorized symptoms as present or absent. Symptoms were coded as present if they appeared in the WeChat group chat texts (eg, menorrhagia, dysmenorrhea, menstrual disorder, or prolonged menstrual period). In contrast, if the symptom appeared in the WeChat group chat texts but included a negation term, it was coded as absent (eg, no dysmenorrhea and no vaginal secretion).

Once symptoms were defined, 2 members of the study team (annotators) reviewed each note in the training data set. After the 2 annotators coded their notes, a third investigator reviewed the annotations and resolved discrepancies. Doccano software was used to annotate the documents. Doccano allows reviewers to view words with coding concordance (highlighted in green) or discordance (highlighted in red). If the first 2 annotators were unsure of how to resolve discordant coding, all 3 study team members resolved the coding through discussion and, when necessary, formalized the iterative coding rules. The frequency distribution of manually coded symptoms is shown in Figure S2 in [Multimedia Appendix 1](#).

## Establishment of Symptomatic Corpus

Manual coding generated an initial corpus of relevant symptomatic key terms or phrases that we used to define the features in our models (Table S1 in [Multimedia Appendix 1](#)). Given that the Doccano tool uses regular expressions to define features in models, we wrote a series of regular expressions using the symptomatic key terms and phrases described above, shortened the list of regular expressions by counting the number of features detected in our training data set, and removed those that did not have any counts [15-18].

By manually coding 30% of the WeChat group chat texts, we obtained a corpus of symptomatic key terms and phrases. Subsequently, we developed a text-processing approach based on Python's regular expression skills. The corpus returned from the manual coding was further processed using a number of string-processing functions that support regular expression-like features. We used these functions to process the WeChat group chat texts to identify the postoperative symptom patterns.

## Extraction of Identity Information

We used regular expressions in Python to extract patients' identity information (patient ID and name) from the WeChat group chats. Duplicates were removed. We used these identities to match the patients in the EHR and extracted disease and treatment information for each patient obtained from the WeChat group chat text.

## Algorithm for Symptom Identification

We applied regular expression based on the corpus, which was attained after manual coding, and then identified the symptoms in the corpus (Table S1 in [Multimedia Appendix 1](#)) from the WeChat group chat texts. Using the parsing capability of regular expression in Python, we systematically searched for symptoms on uncoded WeChat group chat texts. After a symptom was identified with all different key terms or phrases, we moved to

another symptom until key terms or phrases of the corpus have been browsed through; simultaneously, we pinpointed the time when each symptom occurred.

## Data Analysis

Symptoms were summarized as frequencies of presence before and after FUAS separately, and continuous variables were expressed as means and SDs or medians and IQRs. Categorical data were presented as percentages. The *t* test (2-tailed) and Wilcoxon paired rank-sum test were used to compare baseline data between the non-dysmenorrhea and dysmenorrhea groups for continuous variables as appropriate.

The text mining of WeChat group chats was compared with manual review findings (gold standard). The agreement was assessed with  $\kappa$  statistics.  $\kappa$  measures the proportion of agreement between methods after removing any chance agreement. For  $\kappa$ , values of 0.61 to 0.80 are considered good, and values of 0.81 to 1.00 are considered excellent.

We defined the top 10 symptoms presented before surgery as *disease-related symptoms* and symptoms that emerged in the text of WeChat after surgery as *treatment-related symptoms*. Multivariate generalized estimating equation (GEE) models were constructed to identify risk factors for the prevalence of symptoms over time (before FUAS and 1, 2, 3, and 4-6 months after FUAS). Factors included in GEE models were age ( $\leq 40$  years vs  $>40$  years), BMI ( $\leq 23.37$  kg/m<sup>2</sup> vs  $>23.37$  kg/m<sup>2</sup>), pregnancy ( $\leq 2$  vs  $>2$ ), parity ( $\leq 1$  vs  $>1$ ), and abortion ( $\leq 2$  vs  $>2$ ). The effects of risk factors were presented as odds ratios (ORs) with 95% CIs. Using the Bonferroni correction for multiple comparisons of risk factor identification, the statistical significance level was set at the adjusted cutoff of *P* value = .05 / number of risk factors.

We used the GEE model to describe the trajectories of the 9 most common symptoms in the order of prevalence (vaginal secretion, dysmenorrhea, menstrual disorder, prolonged menstrual periods, lower abdominal pain, menstrual blood clots, menorrhagia, dizziness, and fatigue) over the 5 time points. Two piecewise (from before to 1 month after FUAS and from 1 month to 4-6 months after FUAS) random coefficient models were used to analyze trends of symptom presence before and after surgery. All *P* values were 2-tailed, and statistical significance was set at the conventional cutoff of *P* < .05. All data analyses were performed using the statistical software SAS (version 9.4; SAS Institute).

## Results

### Data Extraction

We captured the basic information of the patients' medical records from 20 clinician-patient WeChat groups of 5 attending doctors. After excluding patients with other gynecological diseases, 408 patients with UFs were included in this analysis. A total of 8188 screenshots were obtained. We extracted 939,735 characters after image recognition and free-text cleaning. Finally, the group chat texts of patients with UFs contained 190,000 words. Five time points were extracted, including the preoperative period and 1, 2, 3, and 4-6 months after surgery.

## Demographics

Patient baseline characteristics are presented in [Table 1](#). The mean age was 39.94 (SD 6.81; IQR 35-45) years; the mean BMI was 23.47 (SD 29.37; range 20.20-23.83) kg/m<sup>2</sup>; and the median

number of pregnancies, abortions, Caesarean sections, and parity were 2 (IQR 1-3), 2 (IQR 1-3), 0 (IQR 0-1), and 1 (IQR 1-1), respectively. Preoperative symptoms were present in 59.1% (241/408) of patients.

**Table 1.** Demographic characteristics of the study population.

Demographic characteristics	Value (N=408)
Age (years), mean (SD; IQR)	39.94 (6.81; 35-45)
BMI (kg/m <sup>2</sup> ; n=402), mean (SD; IQR)	23.47 (29.37; 20.20-23.83)
Pregnancy (n=406), mean (SD; IQR)	2.36 (1.87; 1-3)
Abortion (n=323), mean (SD; IQR)	1.90 (1.43; 1-3)
Caesarean section (n=290), mean (SD; IQR)	0.46 (0.60; 0-1)
Parity (n=326), mean (SD; IQR)	0.98 (0.63; 1-1)
Age of menarche (years; n=405), mean (SD; IQR)	13.29 (1.51; 12-14)
Menstrual period (days; n=407), mean (SD; IQR)	6.13 (2.59; 5-7)
Menstrual cycle (days; n=405), mean (SD; IQR)	28.76 (4.07; 27-30)
UF <sup>a</sup> -LR <sup>b</sup> size (mm; n=161), mean (SD; IQR)	53.8 (17.7; 45-61)
UF-FB <sup>c</sup> size (mm; n=381), mean (SD; IQR)	52.7 (17.9; 40-63)
Length of UF (mm; n=399), mean (SD; IQR)	60.1 (22.0; 45-70)
Length of stay (days; n=400), mean (SD; IQR)	3.10 (1.97; 2-3)
<b>Readmission, n (%)</b>	
Yes	44 (10.8)
No	364 (89.2)
<b>Location of UF (n=387), n (%)</b>	
Submucosal	22 (5.7)
Intramural	347 (89.7)
Subserosal	18 (4.6)
<b>Employment status (n=407), n (%)</b>	
Unemployed	167 (41)
Part-time employment	1 (0.2)
Full-time employment	239 (58.7)
<b>Ethnicity (n=407), n (%)</b>	
1	389 (95.6)
6	2 (0.5)
7	16 (3.9)
<b>Number of UFs (n=404), n (%)</b>	
Single	140 (34.6)
Multiple	264 (65.4)
<b>Marriage status, n (%)</b>	
1	321 (78.7)
2	34 (8.3)
3	1 (0.2)
5	50 (12.2)
6	2 (0.5)
<b>Extent of dysmenorrhea, n (%)</b>	
None	281 (68.9)
Mild	89 (21.8)
Moderate	29 (7.1)

Demographic characteristics	Value (N=408)
Severe	9 (2.2)

<sup>a</sup>UF: uterine fibroid.

<sup>b</sup>LR: left to right.

<sup>c</sup>FB: front to back.

### Corpus of UFs

According to the manual coding of the concept of postoperative symptoms, the most reported symptomatic concepts were menorrhagia, dysmenorrhea, prolonged menstrual periods, menstrual blood clots, dizziness, fatigue, lower abdominal pain, menstrual disorder, and vaginal secretion (eg, “increased secretion,” “watery secretion,” “bloody secretions,” “excreted secretions,” and “vaginal secretion”). Tables S2 and S3 in [Multimedia Appendix 1](#) present the symptom normalization mapping [7-10].

We extracted the frequency of symptoms before and after FUAS ([Multimedia Appendix 2](#)). The presence rates were significantly

reduced after surgery for several symptoms, including menorrhagia (137/408, 33.6% vs 17/408, 4.2%), dysmenorrhea (135/408, 33.1% vs 64/408, 15.7%), prolonged menstrual period (86/408, 21.1% vs 34/408, 8.3%), menstrual blood clots (55/408, 13.7% vs 28/408, 6.9%), dizziness (55/408, 13.7% vs 10/408, 2.5%), and fatigue (55/408, 13.7% vs 10/408, 2.5%; all  $P < .001$ ). Two symptoms (vaginal secretion [0/408, 0% vs 64/408, 15.7%] and menstrual disorder [16/408, 3.9% vs 62/408, 15.2%]) were reported significantly more frequently after surgery than before FUAS (all  $P < .001$ ). No statistical difference was found in lower abdominal pain before and after FUAS (40/408, 9.8% vs 33/408, 8.1%;  $P = .39$ ; [Table 2](#)).

**Table 2.** Comparison of the prevalence of preoperative and postoperative symptoms (N=408).

Symptoms	Preoperative prevalence, n (%)	Postoperative prevalence, n (%)	$\chi^2$ (df=1)	P value
Vaginal secretion	0 (0)	64 (15.7)	69.447	<.001
Dysmenorrhea	135 (33.1)	64 (15.7)	33.502	<.001
Menstrual disorder	16 (3.9)	62 (15.2)	29.995	<.001
Prolonged menstrual periods	86 (21.1)	34 (8.3)	26.418	<.001
Lower abdominal pain	40 (9.8)	33 (8.1)	0.737	.39
Menorrhagia	137 (33.6)	17 (4.2)	115.259	<.001
Menstrual blood clots	55 (13.7)	28 (6.9)	10.404	<.001
Dizziness	55 (13.7)	10 (2.5)	34.882	<.001
Fatigue	55 (13.7)	10 (2.5)	34.882	<.001

### Symptom Changing Over Time

The trajectories of 10 symptoms were analyzed. Menorrhagia, prolonged menstrual periods, menstrual blood clots, dizziness,

lower abdominal pain, intensified dysmenorrhea, and frequent urination significantly decreased from the preoperative period to 4-6 months after FUAS treatment (all  $P < .001$ ; [Table 3](#)).

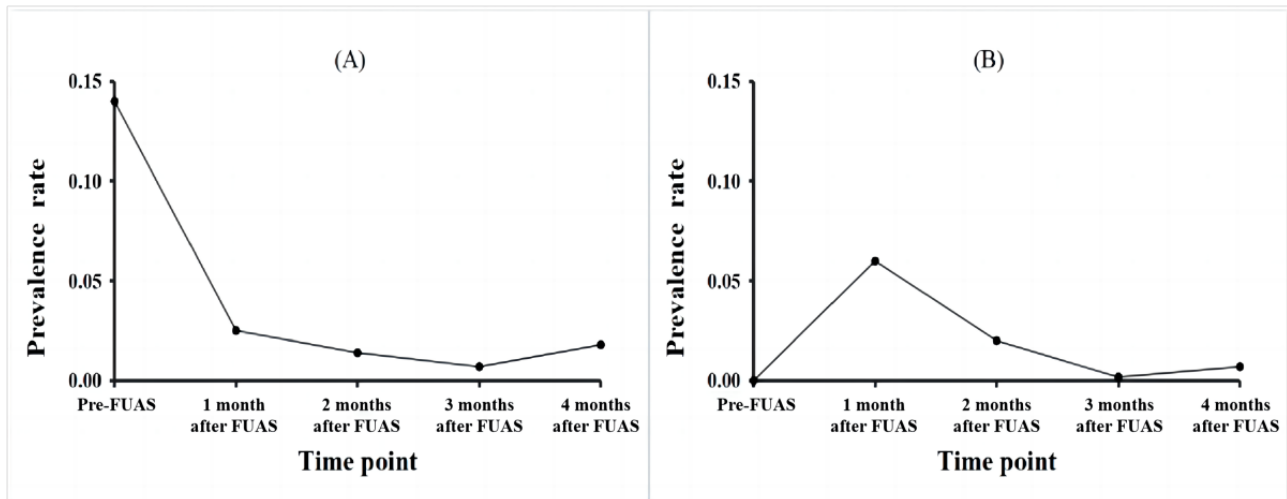
**Table 3.** Generalized estimating equation model analysis of perioperative symptoms at 5 time points.

Variables	Estimate (time; SE)	P value
Menorrhagia	-1.4036 (0.1892)	<.001
Dysmenorrhea	-0.2571 (0.1492)	.08
Prolonged menstrual period	-3.7684 (0.9710)	<.001
Menstrual blood clots	-1.2824 (0.2552)	<.001
Dizziness	-1.8632 (0.4329)	<.001
Lower abdominal pain	-0.6222 (0.1309)	<.001
Menstrual disorder	0.0327 (0.0991)	.74
Vaginal secretion	-0.1359 (0.0754)	.07
Intensified dysmenorrhea	-2.3502 (0.7010)	<.001
Frequent urination	-2.1042 (0.4893)	<.001

Two symptoms demonstrated significantly different trends of change over the 2 time period (from before to 1 month after FUAS and from 1 to 4 months after FUAS). The proportion of patients with dysmenorrhea tended to significantly decrease from the preoperative period to the first postoperative month (estimate  $-0.49$ ;  $P=.03$ ) and gradually stabilized after the first

postoperative month (estimate  $0.17$ ;  $P=.16$ ). The proportion of patients with vaginal secretion significantly increased from the preoperative period, peaking in the first postoperative month (estimate  $1.43$ ;  $P=.04$ ), and decreased thereafter (estimate  $-1.08$ ;  $P=.04$ ; [Figure 1](#)).

**Figure 1.** Two piecewise regression analysis for 2 symptoms during 5 time points in the perioperative period: (A) dysmenorrhea and (B) vaginal secretion. FUAS: focused ultrasound ablation surgery.



### Association Between Preoperative Dysmenorrhea and Postoperative Symptoms

Compared to those without preoperative dysmenorrhea, patients with dysmenorrhea had a higher probability of reporting menstrual blood clots after FUAS (11.11% vs 4.76%,  $P=.02$ ),

whereas the differences in vaginal discharge, menorrhagia, prolonged menstrual period, advanced or delayed menstruation, dysmenorrhea, dizziness, lumbar and abdominal discomfort, and frequent urination were not significant (all  $P<.05$ ; [Table 4](#)).

**Table 4.** Analysis of the difference between preoperative dysmenorrhea and postoperative symptoms.

Symptom and preoperative dysmenorrhea	Postoperative symptoms, n (%)	Statistics	P value
<b>Vaginal secretion</b>		-0.3169	.96
Yes (n=135)	21 (15.6)		
No (n=273)	43 (15.8)		
<b>Menstrual blood clots</b>		11.6663	.02
Yes (n=135)	15 (11.1)		
No (n=273)	13 (4.8)		
<b>Menorrhagia</b>		8.952	.12
Yes (n=135)	21 (15.6)		
No (n=273)	28 (10.3)		
<b>Prolonged menstrual period</b>			
Yes (n=135)	9 (6.7)	-4.4473	.39
No (n=273)	25 (9.2)		
<b>Menstrual disorder</b>		3.2851	.59
Yes (n=135)	22 (16.3)		
No (n=273)	38 (13.9)		
<b>Dysmenorrhea</b>		8.0487	.20
Yes (n=135)	26 (19.3)		
No (n=273)	39 (14.3)		
<b>Dizziness</b>		1.613	.64
Yes (n=135)	4 (3)		
No (n=273)	6 (2.2)		
<b>Lower abdominal pain</b>		8.1021	.12
Yes (n=135)	15 (11.1)		
No (n=273)	18 (6.6)		
<b>Frequent urination</b>		-0.8309	.73
Yes (n=135)	1 (0.7)		
No (n=273)	3 (0.01)		

## Factors Related to Symptoms

### Preoperative Factors

As shown in Table 5, differences in age group ( $P=.006$ ), BMI ( $P=.006$ ), and the number of pregnancies ( $P=.04$ ) between

patients with and without dysmenorrhea were statistically significant in the patient population before FUAS. Patients aged  $\leq 40$  years (OR 0.55, 95% CI 0.36-0.84), those with a BMI greater than  $24.0 \text{ kg/m}^2$  (OR 1.88, 95% CI 1.20-2.94), or those with  $< 2$  pregnancies (OR 0.63, 95% CI 0.41-0.97) were more likely to have dysmenorrhea.



**Table 5.** Risk factor analysis for patients with preoperative dysmenorrhea.

Variables	Patients with dysmenorrhea (n=135), n (%)	Patients without dysmenorrhea (n=262), n (%)	P value
<b>Age group (years)</b>			.006
≤40	77 (57)	111 (42.4)	
>40	58 (43)	151 (57.6)	
<b>BMI (kg/m<sup>2</sup>)</b>			.006
<24.0	84 (62.2)	198 (75.6)	
≥24.0	51 (37.8)	64 (24.4)	
<b>Pregnancy</b>			.04
<2	60 (44.4)	88 (33.6)	
≥2	75 (55.6)	174 (66.4)	
<b>Abortion</b>			.09
<2	79 (58.5)	144 (55)	
≥2	56 (41.5)	118 (45)	
<b>Caesarean section</b>			.35
0	44 (32.6)	122 (46.6)	
>0	37 (27.4)	78 (29.8)	
<b>Parity</b>			.75
≤2	56 (41.5)	85 (32.4)	
>1	79 (58.5)	177 (67.6)	
<b>Age of menophania</b>			.28
<13	83 (61.5)	145 (55.3)	
≥13	52 (38.5)	117 (44.7)	
<b>Menstrual period (days)</b>			.92
<6	76 (56.3)	149 (56.9)	
≥6	59 (43.7)	113 (43.1)	
<b>Menstrual cycle (days)</b>			.67
<29	70 (51.9)	129 (49.2)	
≥29	65 (48.1)	133 (50.8)	
<b>UF<sup>a</sup>-LR<sup>b</sup> size (mm)</b>			.34
<53.8	114 (84.4)	210 (80.2)	
≥53.8	21 (15.6)	52 (19.8)	
<b>Anteroposterior diameter of UF (mm)</b>			.11
<52.7	83 (61.5)	138 (52.7)	
≥52.7	52 (38.5)	124 (47.3)	
<b>Length of UF (mm)</b>			.60
<60.1	75 (55.6)	137 (52.3)	
≥60.1	60 (44.4)	125 (47.7)	
<b>Total energy (KJ)</b>			.29
<398	91 (67.4)	191 (72.9)	
≥398	44 (32.6)	71 (27.1)	
<b>Total irradiation time (s)</b>			.75
<89.1	74 (54.8)	149 (56.9)	

Variables	Patients with dysmenorrhea (n=135), n (%)	Patients without dysmenorrhea (n=262), n (%)	P value
≥89.1	61 (45.2)	113 (43.1)	
<b>Total treatment time (m)</b>			.27
<895.9	82 (60.7)	174 (66.4)	
≥895.9	53 (39.3)	88 (33.6)	
<b>Position of UF</b>			.51
Submucosal	5 (3.7)	17 (6.5)	
Intramural	118 (87.4)	219 (83.6)	
Subserosal	5 (3.7)	13 (5)	
<b>Number of UFs</b>			.44
Single	47 (34.8)	89 (34)	
Multiple	86 (63.7)	171 (65.3)	

<sup>a</sup>UF: uterine fibroid.

<sup>b</sup>LR: left to right.

### Postoperative Factors

As shown in Table 6, the risk factor for postoperative lower abdominal pain was the length of fibroids, and the results were statistically significant ( $P=.03$ ). The risk factor for menorrhagia

in postoperative patients was the age at menarche, and the results were statistically significant ( $P=.02$ ). The risk factor for postoperative menstrual blood clots was the number of fibroids, and the results were statistically significant ( $P=.04$ ).

**Table 6.** Risk factors analysis of postoperative major symptomatic concepts.

Variables	Risk factor	Estimates (SE)	P value
Vaginal secretion	Menstrual period (days)	1.1421 (0.4358)	-.009
Dysmenorrhea disorder	None	— <sup>a</sup>	—
Menstrual disorder	None	—	—
Prolonged menstrual period	None	—	—
Lower abdominal pain	Length of UFs <sup>b</sup>	0.1190 (0.0563)	.03
Menorrhagia	Age of menophania	-1.0249 (0.4382)	.02
Menstrual blood clots	Number of UFs	2.3357 (1.1269)	.04
Dizziness	None	—	—
Fatigue	None	—	—

<sup>a</sup>Not applicable.

<sup>b</sup>UF: uterine fibroid.

## Discussion

### Principal Findings

For the first time, we extracted the concepts of patients' symptoms from unstructured WeChat group chats text, demonstrating the feasibility of capturing what patients experienced after treatment from the most popular social media platform, WeChat, which features private communication between patients and medical staff. By combining medical records before surgery and WeChat communication after surgery, we profiled the patients' experiences with UFs over the full course of surgical treatment. By establishing a relevant symptom corpus, we provided a research foundation to develop

specific instruments for the extensive care of patients with UFs after surgery [6,19,20].

UFs are one of the most common benign tumors in women, although their prevalence may be underestimated in asymptomatic women. Approximately 25% to 50% of women with fibroids are symptomatic and experience heavy menses, reproductive issues, pain, increased urinary frequency, and anemia. Treatments for UFs, whether invasive, minimally invasive, or noninvasive, relieve disease-related symptoms and induce symptoms due to surgical insults [21]. Accurate monitoring of symptom trajectories over the course of surgery and recovery would help evaluate the treatment effect, manage postoperative complications, and alert about disease recurrence [22].

Using the real-world data from clinician-patient WeChat communication and the EHR, we identified 2 categories of symptoms, including disease-related symptoms and treatment-related symptoms. In this study, the top 5 symptoms collected during the postoperative rehabilitation of FUAS were vaginal secretion, dysmenorrhea, menstrual disorder, prolonged menstrual period, and lower abdominal pain. The symptoms involved in our study were similar to those in the study by Jeng et al [22], in which patients reported symptoms of UFs and health-related quality of life (Uterine Fibroid Symptom and Health-Related Quality of Life [UFS-QOL] questionnaire [23]). The application of a comprehensive questionnaire, such as the UFS-QOL, in clinical research, standardizes the outcome assessment but requires time and personnel resources, which may not be available in real-world practice for patient care. Moreover, although symptoms can be assessed using a disease-specific questionnaire, as a newly developed noninvasive surgical technique, FUAS-related symptoms have not been well validated in any PRO tool. Thus, the concept of symptoms in patients undergoing FUAS should first be based on qualitative studies.

Our study used NLP to identify symptoms experienced by patients after FUAS in WeChat group communication, providing a list of items that may serve as a draft of a FUAS-specific PRO symptom assessment tool. Although WeChat group communication may not produce a clear theme [24,25], it saves time and expenses for interviewer training, recruitment, and interviews required in a traditional qualitative study. Moreover, qualitative interviews are influenced by the current situation of patients [26], whereas the WeChat group communication provided the contemporary record of patients' status and needs for health care during a period of time. In addition, compared with the small number of patients in a qualitative interview, the sample size of data from clinician-patient group communication could be large because of the popularity of WeChat in China. Thus, using WeChat data might improve the representativeness of patient experiences and could be representative of a more generalized population.

Social media communication may also inform patient management in the hospital [27,28]. Existing medical resources cannot fulfill patients' needs on time, and it is also costly to consult outpatient services [29]. When minor problems emerge, people often turn to social media platforms to identify patients with similar experiences or web-based groups of doctors for advice. Reporting free texts is often unstructured but has potential. As patients actively report their health conditions, capturing adverse clinical events from this specific source is possible [30]. Conversely, medical staff can track patients' health status and indicators on the web, collect relevant health data, and conduct appropriate scientific research activities.

## Acknowledgments

The authors thank Dan Yu (Chongqing Haifu Hospital) for her contribution to the data collection of this paper. We also thank Jianhao Chen for computer programming support.

This research was funded by The Chongqing Postgraduate Scientific Research Innovation Project (CYB22220) and The National Key Research and Development Program of the Ministry of Science and Technology of China—Intergovernmental International Cooperation Project on Science and Technology Innovation (2022YFE0133100).

One hurdle in using social media data is that text is unstructured. NLP technology has been described in the literature as an effective, inexpensive, and highly accurate tool for automatically identifying concepts from unstructured patient text [31,32]. NLP technology enables the accommodation of large unstructured data, significantly reducing time and labor costs that are spent in a traditional qualitative interview. It can quickly provide a framework for conventional qualitative analysis and then add patients in qualitative interviews to saturate it. Further research should be conducted on how NLP technology can be combined with qualitative interviews for the purpose of generating clinically meaningful results from unstructured medical documentations [24,25,33,34]

Our study achieved a high agreement between manual coding and Python-based text mining, similar to other studies that used SAS-based text mining [35-37]. This consistency suggests the potential of accurately extracting symptoms from free texts without the need for a lengthy, comprehensive manual review process. The process of NLP is not discourse or population specific, including compiling lists of search terms; manually reviewing samples of the data set to investigate abbreviations, misspellings, and negations; determining the order of search groups and elimination records; and investigating matches and mismatches throughout the process. This concept can be applied to other studies that investigate treatments in free texts.

## Limitations

First, the median period for data collection ended at 4-6 months after treatment, which is not sufficient for recurrence identification; this is a major concern in the management of UFs. Second, the symptoms identified from WeChat and the EHR were not scored for severity. Using the obtained list of symptoms, instruments with an appropriate scale for each symptom can be developed and validated for the precise evaluation of patients' status. Third, the patients in this study were from a private hospital with a higher socioeconomic status than patients with UFs in the general population [1]. For broader interpretation, the current results should be validated in public hospitals in different areas.

## Conclusions

Our study showed that the extraction of a large amount of health data from social media is feasible. Python-based text mining is accurate and efficient in summarizing the symptoms experienced by patients with UFs under FUAS treatment. Mining algorithms provide a basis for the rapid and efficient analysis of large samples of free text in the future. The symptom burden of UFs treated with FUAS as summarized in this study can also provide a relevant reference for subsequent rehabilitation of patients.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Supplemental materials.

[\[DOCX File , 330 KB-Multimedia Appendix 1\]](#)

## Multimedia Appendix 2

Distribution of the most common symptoms per category.

[\[DOCX File , 333 KB-Multimedia Appendix 2\]](#)

## References

1. Guntuku SC, Sherman G, Stokes DC, Agarwal AK, Seltzer E, Merchant RM, et al. Tracking mental health and symptom mentions on Twitter during COVID-19. *J Gen Intern Med* 2020 Sep;35(9):2798-2800 [FREE Full text] [doi: [10.1007/s11606-020-05988-8](https://doi.org/10.1007/s11606-020-05988-8)] [Medline: [32638321](https://pubmed.ncbi.nlm.nih.gov/32638321/)]
2. Wakamiya S, Morita M, Kano Y, Ohkuma T, Aramaki E. Tweet classification toward Twitter-based disease surveillance: new data, methods, and evaluations. *J Med Internet Res* 2019 Feb 20;21(2):e12783 [FREE Full text] [doi: [10.2196/12783](https://doi.org/10.2196/12783)] [Medline: [30785407](https://pubmed.ncbi.nlm.nih.gov/30785407/)]
3. Mowery D, Smith H, Cheney T, Stoddard G, Coppersmith G, Bryan C, et al. Understanding depressive symptoms and psychosocial stressors on Twitter: a corpus-based study. *J Med Internet Res* 2017 Feb 28;19(2):e48 [FREE Full text] [doi: [10.2196/jmir.6895](https://doi.org/10.2196/jmir.6895)] [Medline: [28246066](https://pubmed.ncbi.nlm.nih.gov/28246066/)]
4. Yom-Tov E, Borsa D, Cox IJ, McKendry RA. Detecting disease outbreaks in mass gatherings using internet data. *J Med Internet Res* 2014 Jun 18;16(6):e154 [FREE Full text] [doi: [10.2196/jmir.3156](https://doi.org/10.2196/jmir.3156)] [Medline: [24943128](https://pubmed.ncbi.nlm.nih.gov/24943128/)]
5. Stewart EA, Laughlin-Tommaso SK, Catherino WH, Lalitkumar S, Gupta D, Vollenhoven B. Uterine fibroids. *Nat Rev Dis Primers* 2016 Jun 23;2:16043 [doi: [10.1038/nrdp.2016.43](https://doi.org/10.1038/nrdp.2016.43)] [Medline: [27335259](https://pubmed.ncbi.nlm.nih.gov/27335259/)]
6. Donnez J, Dolmans MM. Uterine fibroid management: from the present to the future. *Hum Reprod Update* 2016 Nov;22(6):665-686 [FREE Full text] [doi: [10.1093/humupd/dmw023](https://doi.org/10.1093/humupd/dmw023)] [Medline: [27466209](https://pubmed.ncbi.nlm.nih.gov/27466209/)]
7. Ning G, Zhang X, Zhang Q, Wang Z, Liao H. Real-time and multimodality image-guided intelligent HIFU therapy for uterine fibroid. *Theranostics* 2020 Mar 26;10(10):4676-4693 [FREE Full text] [doi: [10.7150/thno.42830](https://doi.org/10.7150/thno.42830)] [Medline: [32292522](https://pubmed.ncbi.nlm.nih.gov/32292522/)]
8. Verpalen IM, Anneveldt KJ, Nijholt IM, Schutte JM, Dijkstra JR, Franx A, et al. Magnetic resonance-high intensity focused ultrasound (MR-HIFU) therapy of symptomatic uterine fibroids with unrestrictive treatment protocols: a systematic review and meta-analysis. *Eur J Radiol* 2019 Nov;120:108700 [doi: [10.1016/j.ejrad.2019.108700](https://doi.org/10.1016/j.ejrad.2019.108700)] [Medline: [31634683](https://pubmed.ncbi.nlm.nih.gov/31634683/)]
9. Wang Y, Geng J, Bao H, Dong J, Shi J, Xi Q. Comparative effectiveness and safety of high-intensity focused ultrasound for uterine fibroids: a systematic review and meta-analysis. *Front Oncol* 2021 Mar 9;11:600800 [FREE Full text] [doi: [10.3389/fonc.2021.600800](https://doi.org/10.3389/fonc.2021.600800)] [Medline: [33767979](https://pubmed.ncbi.nlm.nih.gov/33767979/)]
10. Ou K, Jeng C, Long C, Chuang L. Re: A comparison of the cost-utility of ultrasound-guided high-intensity focused ultrasound and hysterectomy for adenomyosis: a retrospective study: is the cost-effectiveness of HIFU for adenomyosis and fibroids feasible? *BJOG* 2018 May;125(6):763-764 [doi: [10.1111/1471-0528.15115](https://doi.org/10.1111/1471-0528.15115)] [Medline: [29405542](https://pubmed.ncbi.nlm.nih.gov/29405542/)]
11. Liu Y, Ran W, Shen Y, Feng W, Yi J. High-intensity focused ultrasound and laparoscopic myomectomy in the treatment of uterine fibroids: a comparative study. *BJOG* 2017 Aug;124 Suppl 3:36-39 [doi: [10.1111/1471-0528.14745](https://doi.org/10.1111/1471-0528.14745)] [Medline: [28856854](https://pubmed.ncbi.nlm.nih.gov/28856854/)]
12. Tao C, Filannino M, Uzuner. Prescription extraction using CRFs and word embeddings. *J Biomed Inform* 2017 Aug;72:60-66 [FREE Full text] [doi: [10.1016/j.jbi.2017.07.002](https://doi.org/10.1016/j.jbi.2017.07.002)] [Medline: [28684255](https://pubmed.ncbi.nlm.nih.gov/28684255/)]
13. Peng J, Zhao M, Havrilla J, Liu C, Weng C, Guthrie W, et al. Natural language processing (NLP) tools in extracting biomedical concepts from research articles: a case study on autism spectrum disorder. *BMC Med Inform Decis Mak* 2020 Dec 30;20(Suppl 11):322 [FREE Full text] [doi: [10.1186/s12911-020-01352-2](https://doi.org/10.1186/s12911-020-01352-2)] [Medline: [33380331](https://pubmed.ncbi.nlm.nih.gov/33380331/)]
14. Manyonda I, Belli A, Lumsden M, Moss J, McKinnon W, Middleton LJ, FEMME Collaborative Group. Uterine-artery embolization or myomectomy for uterine fibroids. *N Engl J Med* 2020 Jul 30;383(5):440-451 [doi: [10.1056/NEJMoa1914735](https://doi.org/10.1056/NEJMoa1914735)] [Medline: [32726530](https://pubmed.ncbi.nlm.nih.gov/32726530/)]
15. Liu H, Li Q, Zhan Y, Zhang Z, Zeng DD, Leischow SJ. Characterizing social media messages related to underage JUUL e-cigarette buying and selling: cross-sectional analysis of Reddit subreddits. *J Med Internet Res* 2020 Jul 20;22(7):e16962 [FREE Full text] [doi: [10.2196/16962](https://doi.org/10.2196/16962)] [Medline: [32706661](https://pubmed.ncbi.nlm.nih.gov/32706661/)]
16. Frenz CM. Deafness mutation mining using regular expression based pattern matching. *BMC Med Inform Decis Mak* 2007 Oct 25;7:32 [FREE Full text] [doi: [10.1186/1472-6947-7-32](https://doi.org/10.1186/1472-6947-7-32)] [Medline: [17961241](https://pubmed.ncbi.nlm.nih.gov/17961241/)]

17. Ciofi Degli Atti ML, Pecoraro F, Piga S, Luzi D, Raponi M. Developing a surgical site infection surveillance system based on hospital unstructured clinical notes and text mining. *Surg Infect (Larchmt)* 2020 Oct;21(8):716-721 [doi: [10.1089/sur.2019.238](https://doi.org/10.1089/sur.2019.238)] [Medline: [32105569](https://pubmed.ncbi.nlm.nih.gov/32105569/)]
18. Turchin A, Kolatkar NS, Grant RW, Makhni EC, Pendergrass ML, Einbinder JS. Using regular expressions to abstract blood pressure and treatment intensification information from the text of physician notes. *J Am Med Inform Assoc* 2006 Nov 1;13(6):691-695 [FREE Full text] [doi: [10.1197/jamia.M2078](https://doi.org/10.1197/jamia.M2078)] [Medline: [16929043](https://pubmed.ncbi.nlm.nih.gov/16929043/)]
19. de la Cruz MSD, Buchanan EM. Uterine fibroids: diagnosis and treatment. *Am Fam Physician* 2017 Jan 15;95(2):100-107 [FREE Full text] [Medline: [28084714](https://pubmed.ncbi.nlm.nih.gov/28084714/)]
20. Martínez-deMiguel C, Segura-Bedmar I, Chacón-Solano E, Guerrero-Aspizua S. The RareDis corpus: a corpus annotated with rare diseases, their signs and symptoms. *J Biomed Inform* 2022 Jan;125:103961 [FREE Full text] [doi: [10.1016/j.jbi.2021.103961](https://doi.org/10.1016/j.jbi.2021.103961)] [Medline: [34879250](https://pubmed.ncbi.nlm.nih.gov/34879250/)]
21. Hsieh C, Lu Y, Liang S. The health concerns related to hysterectomy among Taiwanese women with uterine fibroids- a mixed-methods study. *Women Health* 2021 Jul;61(6):581-590 [doi: [10.1080/03630242.2021.1938791](https://doi.org/10.1080/03630242.2021.1938791)] [Medline: [34187330](https://pubmed.ncbi.nlm.nih.gov/34187330/)]
22. Jeng C, Ou K, Long C, Chuang L, Ker C. 500 cases of high-intensity focused ultrasound (HIFU) ablated uterine fibroids and adenomyosis. *Taiwan J Obstet Gynecol* 2020 Nov;59(6):865-871 [FREE Full text] [doi: [10.1016/j.tjog.2020.09.013](https://doi.org/10.1016/j.tjog.2020.09.013)] [Medline: [33218403](https://pubmed.ncbi.nlm.nih.gov/33218403/)]
23. Yeung S, Kwok JW, Law S, Chung JP, Chan SS. Uterine Fibroid Symptom and Health-related Quality of Life Questionnaire: a Chinese translation and validation study. *Hong Kong Med J* 2019 Dec 4;25(6):453-459 [FREE Full text] [doi: [10.12809/hkmj198064](https://doi.org/10.12809/hkmj198064)] [Medline: [31796639](https://pubmed.ncbi.nlm.nih.gov/31796639/)]
24. Guetterman TC, Chang T, DeJonckheere M, Basu T, Scruggs E, Vydiswaran VGV. Augmenting qualitative text analysis with natural language processing: methodological study. *J Med Internet Res* 2018 Jun 29;20(6):e231 [FREE Full text] [doi: [10.2196/jmir.9702](https://doi.org/10.2196/jmir.9702)] [Medline: [29959110](https://pubmed.ncbi.nlm.nih.gov/29959110/)]
25. Kalvesmaki AF, Chapman AB, Peterson KS, Pugh MJ, Jones M, Gleason TC. Analysis of a national response to a White House directive for ending veteran suicide. *Health Serv Res* 2022 Jun;57 Suppl 1(Suppl 1):32-41 [FREE Full text] [doi: [10.1111/1475-6773.13931](https://doi.org/10.1111/1475-6773.13931)] [Medline: [35238027](https://pubmed.ncbi.nlm.nih.gov/35238027/)]
26. DiMartino L, Miano T, Wessell K, Bohac B, Hanson LC. Identification of uncontrolled symptoms in cancer patients using natural language processing. *J Pain Symptom Manage* 2022 Apr;63(4):610-617 [FREE Full text] [doi: [10.1016/j.jpainsymman.2021.10.014](https://doi.org/10.1016/j.jpainsymman.2021.10.014)] [Medline: [34743011](https://pubmed.ncbi.nlm.nih.gov/34743011/)]
27. Tang SYQ, Israel JS, Poore SO, Affi AM. Facebook facts: breast reconstruction patient-reported outcomes using social media. *Plast Reconstr Surg* 2018 May;141(5):1106-1113 [doi: [10.1097/PRS.0000000000004275](https://doi.org/10.1097/PRS.0000000000004275)] [Medline: [29697604](https://pubmed.ncbi.nlm.nih.gov/29697604/)]
28. Blumenthal KG, Topaz M, Zhou L, Harkness T, Sa'adon R, Bar-Bachar O, et al. Mining social media data to assess the risk of skin and soft tissue infections from allergen immunotherapy. *J Allergy Clin Immunol* 2019 Jul;144(1):129-134 [FREE Full text] [doi: [10.1016/j.jaci.2019.01.029](https://doi.org/10.1016/j.jaci.2019.01.029)] [Medline: [30721764](https://pubmed.ncbi.nlm.nih.gov/30721764/)]
29. Nguyen AX, Trinh X, Wang SY, Wu AY. Determination of patient sentiment and emotion in ophthalmology: infoveillance tutorial on web-based health forum discussions. *J Med Internet Res* 2021 May 17;23(5):e20803 [FREE Full text] [doi: [10.2196/20803](https://doi.org/10.2196/20803)] [Medline: [33999001](https://pubmed.ncbi.nlm.nih.gov/33999001/)]
30. Stokes DC, Andy A, Guntuku SC, Ungar LH, Merchant RM. Public priorities and concerns regarding COVID-19 in an online discussion forum: longitudinal topic modeling. *J Gen Intern Med* 2020 Jul 12;35(7):2244-2247 [FREE Full text] [doi: [10.1007/s11606-020-05889-w](https://doi.org/10.1007/s11606-020-05889-w)] [Medline: [32399912](https://pubmed.ncbi.nlm.nih.gov/32399912/)]
31. Li K, Zhou C, Luo XR, Benitez J, Liao Q. Impact of information timeliness and richness on public engagement on social media during COVID-19 pandemic: an empirical investigation based on NLP and machine learning. *Decis Support Syst* 2022 Nov;162:113752 [FREE Full text] [doi: [10.1016/j.dss.2022.113752](https://doi.org/10.1016/j.dss.2022.113752)] [Medline: [35185227](https://pubmed.ncbi.nlm.nih.gov/35185227/)]
32. Wang Y, Zhang Y, Bennell K, White DK, Wei J, Wu Z, et al. Physical distancing measures and walking activity in middle-aged and older residents in Changsha, China, during the COVID-19 epidemic period: longitudinal observational study. *J Med Internet Res* 2020 Oct 26;22(10):e21632 [FREE Full text] [doi: [10.2196/21632](https://doi.org/10.2196/21632)] [Medline: [33027035](https://pubmed.ncbi.nlm.nih.gov/33027035/)]
33. Deleger L, Molnar K, Savova G, Xia F, Lingren T, Li Q, et al. Large-scale evaluation of automated clinical note de-identification and its impact on information extraction. *J Am Med Inform Assoc* 2013 Jan 01;20(1):84-94 [FREE Full text] [doi: [10.1136/amiajnl-2012-001012](https://doi.org/10.1136/amiajnl-2012-001012)] [Medline: [22859645](https://pubmed.ncbi.nlm.nih.gov/22859645/)]
34. Mellia JA, Basta MN, Toyoda Y, Othman S, Elfanagely O, Morris MP, et al. Natural language processing in surgery: a systematic review and meta-analysis. *Ann Surg* 2021 May 01;273(5):900-908 [doi: [10.1097/SLA.0000000000004419](https://doi.org/10.1097/SLA.0000000000004419)] [Medline: [33074901](https://pubmed.ncbi.nlm.nih.gov/33074901/)]
35. Carrell DS, Halgrim S, Tran D, Buist D, Chubak J, Chapman W, et al. Using natural language processing to improve efficiency of manual chart abstraction in research: the case of breast cancer recurrence. *Am J Epidemiol* 2014 Mar 15;179(6):749-758 [FREE Full text] [doi: [10.1093/aje/kwt441](https://doi.org/10.1093/aje/kwt441)] [Medline: [24488511](https://pubmed.ncbi.nlm.nih.gov/24488511/)]
36. Osborne JD, Wyatt M, Westfall AO, Willig J, Bethard S, Gordon G. Efficient identification of nationally mandated reportable cancer cases using natural language processing and machine learning. *J Am Med Inform Assoc* 2016 Nov;23(6):1077-1084 [FREE Full text] [doi: [10.1093/jamia/ocw006](https://doi.org/10.1093/jamia/ocw006)] [Medline: [27026618](https://pubmed.ncbi.nlm.nih.gov/27026618/)]

37. Xie F, Lee J, Munoz-Plaza CE, Hahn EE, Chen W. Application of text information extraction system for real-time cancer case identification in an integrated healthcare organization. *J Pathol Inform* 2017 Jan;8(1):48 [FREE Full text] [doi: [10.4103/jpi.jpi\\_55\\_17](https://doi.org/10.4103/jpi.jpi_55_17)] [Medline: [29416911](https://pubmed.ncbi.nlm.nih.gov/29416911/)]

## Abbreviations

**EHR:** electronic health record

**FUAS:** focused ultrasound ablation surgery

**GEE:** generalized estimating equation

**NLP:** natural language processing

**OR:** odds ratio

**PRO:** patient-reported outcome

**UF:** uterine fibroid

**UFS-QOL:** Uterine Fibroid Symptom and Health-Related Quality of Life

*Edited by A Mavragani; submitted 02.11.22; peer-reviewed by Y Wang, C Zeng; comments to author 07.12.22; revised version received 26.12.22; accepted 24.07.23; published 01.09.23*

*Please cite as:*

*Zhang J, Xu W, Lei C, Pu Y, Zhang Y, Zhang J, Yu H, Su X, Huang Y, Gong R, Zhang L, Shi Q*

*Using Clinician-Patient WeChat Group Communication Data to Identify Symptom Burdens in Patients With Uterine Fibroids Under Focused Ultrasound Ablation Surgery Treatment: Qualitative Study*

*JMIR Form Res* 2023;7:e43995

URL: <https://formative.jmir.org/2023/1/e43995>

doi: [10.2196/43995](https://doi.org/10.2196/43995)

PMID: [37656501](https://pubmed.ncbi.nlm.nih.gov/37656501/)

©Jiayuan Zhang, Wei Xu, Cheng Lei, Yang Pu, Yubo Zhang, Jingyu Zhang, Hongfan Yu, Xueyao Su, Yanyan Huang, Ruoyan Gong, Lijun Zhang, Qiuling Shi. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 01.09.2023. 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 Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.