HHS Public Access

Author manuscript

*Am J Obstet Gynecol*. Author manuscript; available in PMC 2018 August 30.

Published in final edited form as:

*Am J Obstet Gynecol*. 2018 June ; 218(6): 610.e1–610.e7. doi:10.1016/j.ajog.2018.02.002.

# Optimizing Research in Symptomatic Uterine Fibroids with Development of a Computable Phenotype for Use with Electronic Health Records

**Sarah R HOFFMAN, MS, MPH**[1], **Anissa I VINES, MS, PhD**[1], **Jacqueline R HALLADAY, MD, MPH**[2], **Emily PFAFF, MSIS**[3], **Lauren SCHIFF, MD**[4], **Daniel WESTREICH, PhD**[1], **Aditi SUNDARESAN, MPH**[4,7], **La-Shell JOHNSON, MA**[4,7], and **Wanda K NICHOLSON, MD, MPH, MBA**[4,5,6,7]

[1]Department of Epidemiology, University of North Carolina (UNC) Gillings School of Global Public Health, Chapel Hill, NC, USA, 27599.

[2]Department of Family Medicine, North Carolina Translational and Clinical Sciences (TraCS) Institute, Chapel Hill, NC, USA, 27599.

[3]North Carolina Translational and Clinical Sciences (TraCS) Institute, Chapel Hill, NC, USA, 27599.

[4]Department of Obstetrics and Gynecology, the Center for Women's Health Research, Chapel Hill, NC, USA, 27599.

[5]the Center for Women's Health Research, Chapel Hill, NC, USA, 27599.

[6]Program on Women's Endocrine and Reproductive Health, UNC School of Medicine and the Center for Health Promotion and Disease Prevention, Chapel Hill, NC, USA, 27599

[7]Center for Health Promotion and Disease Prevention, Chapel Hill, NC, USA, 27599

## Abstract

**Background:** Symptomatic uterine fibroids, due to menorrhagia, pelvic pain, bulk symptoms or infertility, are a source of substantial morbidity for reproductive-age women. Comparing Treatment Options for Uterine Fibroids (COMPARE-UF) is a multi-site registry study to compare

Corresponding Author: Wanda Nicholson, MD, MPH, MBA, Professor, Obstetrics and Gynecology, University of North Carolina, 3027 Old Clinic Building; CB# 7570, Chapel Hill, NC 27599-7570.

Disclosures: The authors report no conflict(s) of interest.

Condensation

Computable phenotyping algorithms may be an efficient approach for recruitment of women with symptomatic fibroids for observational and interventional studies.

Implications and Contributions

Women with symptomatic uterine fibroids can report a myriad of symptoms, including pain, bleeding, infertility and psychosocial sequelae. Optimizing fibroid research requires the ability to enroll populations of women with image-confirmed symptomatic uterine fibroids. Our objective was to develop an electronic health record (EHR)-based algorithm to identify women with symptomatic uterine fibroids for a comparative effectiveness study of medical or surgical treatments on quality of life measures. Using an iterative process and text mining techniques, an effective computable phenotype algorithm, comprised of demographics, clinical and laboratory characteristics, was developed with reasonable performance. Such algorithms provide a feasible, efficient way to identify populations of women with symptomatic uterine fibroids for the conduct of large traditional or pragmatic trials and observational comparative effectiveness studies.

the effectiveness of hormonal or surgical fibroid treatments on women's perceptions of their quality of life. Electronic health record (EHR)-based algorithms are able to identify large numbers of women with fibroids, but additional work is needed to develop EHR algorithms that can identify women with *symptomatic* fibroids to optimize fibroid research.

**Objective:** To develop an efficient electronic health record (EHR)-based algorithm that can identify women with symptomatic uterine fibroids in a large health care system for recruitment into large-scale observational and interventional research in fibroid management.

**Study Design:** We developed and assessed the accuracy of three algorithms to identify patients with symptomatic fibroids using an iterative approach. The data source was the Carolina Data Warehouse for Health, a repository for the health system's EHR data. In addition to ICD-9 diagnosis and procedure codes and clinical characteristics, text data mining software was used to derive information from imaging reports to confirm the presence of uterine fibroids. Results of each algorithm were compared with expert manual review to calculate the positive predictive values (PPVs) for each algorithm.

**Results:** Algorithm 1 was composed of the following criteria: i) age 18–54 AND ii) either ≥ 1 ICD-9 diagnosis codes for uterine fibroids or mention of fibroids using text-mined key words in imaging records or documents AND iii) no ICD-9 or CPT codes for hysterectomy and no reported history of hysterectomy. The PPV was 47% (95% CI: 39%−56%). Algorithm 2 required i) ≥ 2 ICD-9 diagnosis codes for fibroids AND positive text mined key words and had a PPV of 65% (95% CI: 50%−79%). In Algorithm 3, further refinements included ≥ 2 ICD-9 diagnosis codes for fibroids on separate outpatient visit dates, the exclusion of women who had a positive pregnancy test within three months of their fibroid-related visit, and exclusion of incidentally detected fibroids during prenatal or emergency department visits. Algorithm 3 achieved a PPV of 76% (95% CI: 71%−81%).

**Conclusion:** An EHR-based algorithm is capable of identifying cases of symptomatic uterine fibroids with moderate positive predictive value and may be an efficient approach for large-scale study recruitment.

### Keywords

electronic health records; uterine fibroids; validation; computable phenotype; positive predictive value; women's health; comparative effectiveness research

## INTRODUCTION

While electronic health records (EHRs) were launched initially for improvement and transferability of patient care documentation and billing, there has been growing interest from hospital leaders, federal agencies, and researchers to maximize use of EHRs for clinical research. Recent attention by researchers[1,2] and federal funders[3] has focused on the use of EHRs to speed or enhance participant recruitment, particularly for large pragmatic trials that focus on a broad range of diseases or conditions, prospective comparative effectiveness studies, and for rare diseases or clinical conditions for which there is limited evidence on treatment effectiveness.[4]

Uterine fibroids (also referred to as leiomyomas or myomas) are common, benign tumors of the uterus, reported in 60%−80% of reproductive-age women.[5] The majority of women with fibroids are asymptomatic, but about 20% of women can experience menorrhagia, pelvic pain, bulk symptoms (i.e. pelvic pressure or genitourinary symptoms) or infertility.[6] Fibroid symptoms can have a substantial impact on women's daily activities, reducing their quality of life and posing a significant emotional burden.[7,8] Despite the prevalence and burden of uterine fibroids, there is limited population-based research on the comparative effectiveness of medical versus surgical treatments. In response to this knowledge gap of treatment effectiveness, the Agency for Healthcare Research and Quality (AHRQ)[9] and the Patient-Centered Outcomes Research Institute (PCORI) issued a report[10] and priority agenda for comparative effectiveness research in uterine fibroids. Part of this research agenda includes a national registry of women with symptomatic uterine fibroids and the conduct of large, prospective studies to compare patient-centered outcomes among women undergoing medical and surgical treatments.[5] The report also highlighted current limitations in using administrative data to identify symptomatic patients for registry-based fibroid research and women's population health.

The Comparing Treatments Options for Uterine Fibroids (COMPARE-UF) Study[11] is a multi-site initiative (Mayo (Rochester), INOVA Health System, Brigham and Women's, University of Mississippi, University of California Fibroid Network, Henry Ford Health System, University of Michigan, Department of Defense, and University of North Carolina) to develop a national registry of women with symptomatic uterine fibroids and prospectively compare the effectiveness of different medical (i.e. hormonal) and surgical treatments on women's perceptions of their quality of life. Conducting a large registry study of women with uterine fibroids requires an effective EHR algorithm that can be used reliably to identify and classify women with fibroids as symptomatic or asymptomatic for appropriate, efficient, and timely registry recruitment and study enrollment.

While prior studies have developed EHR-based algorithms to identify women with image-confirmed uterine fibroids,[12] there has been little effort to develop a computable EHR phenotype to identify women with *symptomatic* fibroids, as demonstrated by documentation of bleeding, pain, bulk symptoms or infertility. Our objectives were to: 1) develop and refine computable phenotyping algorithms that can efficiently identify symptomatic cases of uterine fibroids and 2) assess the performance of each algorithm, as measured by positive predictive value (PPV), using manual expert chart review as the basis for establishing true positive cases of symptomatic fibroids.[1] Parallel to COMPARE-UF, we classified women with infertility related to fibroids as being symptomatic. While there are studies that focus on conception after fibroid management in women with infertility, there is little data on the comparative effectiveness of different surgical treatments on quality of life measures in this subgroup of women with fibroids.

## MATERIALS AND METHODS

Three EHR-based algorithms were developed and refined to identify women with symptomatic uterine fibroids using data from the University of North Carolina Health Care System (UNCHCS) from April 4, 2014 to October 1, 2015. The UNCHCS is a large, not-

for-profit integrated academic health system in central North Carolina and provides care to a large, diverse population of patients across socioeconomic strata. The Institutional Review Board at the University of North Carolina approved the study. Below, we describe the data sources and summarize the steps taken to develop and assess the performance of each algorithm.

## Data Sources

Two sources of data were utilized from the Carolina Data Warehouse for Health (CDW-H),[13] a central data source containing clinical, laboratory and administrative billing data for patients seen within the UNC Healthcare system. Data from UNC EHRs are transferred to the CDW-H and updated on a nightly basis. We used the following data sources:

1. Clinical EHR data included demographics (i.e. age), patient problem lists, laboratory test results, ultrasound and magnetic resonance imaging reports, outpatient and in-patient visits and the type of outpatient visits (e.g. radiology, medical or emergency department (ED) visits) and medications. The patient problem list included the International Classification of Diseases, Ninth Revision Clinical Modification (ICD-9-CM) codes for current and past conditions,
2. Billing data included ICD-9-CM diagnosis codes for uterine fibroids, ICD-9-CM procedure codes, and Current Procedural Terminology (CPT) codes for outpatient, inpatient, and emergency department visits.

## Study Population and Definition of Symptomatic Fibroids.

We used the same inclusion criteria to define the study population as those used for the COMPARE-UF study: non-pregnant women, ages 18 to 54, with ultrasound or magnetic resonance imaging (MRI) that suggests the presence of one or more fibroids, symptomatic uterine fibroids and no history of hysterectomy. Because the purpose of COMPARE-UF is to assess outcomes after surgical or medical treatment of fibroids, participants were required to have an intact uterus with at least one image-confirmed fibroid. The gold standard diagnostic modality for fibroids is gray-scale ultrasound, followed by MRI.[14]

There are currently no gold standard criteria to define symptomatic fibroids. For this study, we considered a patient as having symptomatic fibroids if there was EHR documentation of the symptoms of bleeding, pain, bulk symptoms or infertility, all of which are consistent with the definition of symptomatic uterine fibroids for inclusion in the COMPARE-UF registry.

## Algorithm Development

Algorithms were developed in an iterative fashion (Figure 1) using a combination of demographic inclusion criteria, ICD-9-CM and CPT codes (Table 1), clinical data, and keywords related to fibroids from clinical notes and imaging reports, to identify women with symptomatic fibroids. For each of the developed algorithms, keyword searches for “positive” or “negative” mention of fibroid-related terms [“fibroid,” “leiomyoma,” or “myomata”] within specific reports were made using text-mining techniques. Python coding was used to

search for negative text terms defined as negators (e.g. “without evidence,” “no evidence”) before each of the key terms.

### Validation

Algorithms were validated in an iterative fashion as they were developed using expert review of randomly selected subsets of charts. Charts were viewed sequentially by a doctoral level researcher (SRH), and an Ob/Gyn clinician researcher with clinical experience in uterine fibroids (WN) reviewer. A third clinician reviewer (JH) was available to render a final decision if the first two reviewers were unable to reach a consensus on symptom status. Women were considered positive for symptomatic fibroids if they met any of the following criteria: (i) clinician documentation in the medical record stating that the patient had “symptomatic fibroids” or “fibroids causing symptoms” or otherwise described the case with definitive language indicating that the patient had fibroid-related symptoms; (ii) current fibroids for which medical treatment was being considered or administered or cases in which surgical management was being planned; (iii) patient had undergone surgical management or additional medical treatment specific to uterine fibroids since the date of the algorithm run.

### Algorithm Performance

There are no established thresholds for evaluating the performance of automated algorithms for use in population-based research, particularly for case-detection of symptomatic uterine fibroids where there is no “gold standard” definition of *symptomatic* fibroid criteria. Our goal was to develop algorithms that were simple, translatable for use in other health care systems, and could yield a maximized positive predictive value (PPV). For the purpose of this study, PPV represents the probability that women identified as having symptomatic uterine fibroids based on the algorithm truly have fibroid-related symptoms as determined by the chart review (considered as the true positive), as shown in the equation below:

$$\text{PPV} = \frac{\text{Patients with symptomatic uterine fibroids by expert chart review}}{\text{Patients identified by EHR-based algorithms}}$$

The Clopper-Pearson method[15] was used to calculate confidence intervals for binomial proportions. Analyses were conducted using SAS 9.3 (SAS Institute Inc., Cary, NC, USA).

## RESULTS

### Algorithms

The initial algorithm (Algorithm1) included the following: (i) ages between 18 and 54 AND (ii) either at least one ICD-9-CM diagnosis code(s) for uterine fibroids OR ultrasound or MRI report with positive mention of fibroids, leiomyoma or myomata using text mining AND (iii) no ICD-9-CM or CPT codes for hysterectomy AND no documented surgical history of hysterectomy (Figure 1). Algorithm 1 identified 4,342 presumptive cases of symptomatic uterine fibroids (Table 2). Review of a subset of 150 EHRs of the presumptive cases yielded 71 true cases and a PPV of 47% (95% Confidence Interval: 0.39–0.56). The primary reason for the misclassifications was the incidental finding of uterine fibroids during

the clinical workup or treatment of other gynecologic conditions (e.g. ovarian cysts, pelvic inflammatory disease), pregnancy, or abdominal pain.

Algorithm 2 (Figure 1) was designed to exclude those patients from Algorithm 1 that were incidental, asymptomatic fibroids. The inclusion criteria were modified to require that patients have (i) 2 or more (rather than 1 or more) ICD-9 diagnosis codes for fibroids AND (ii) an imaging report with evidence of fibroids. We thought that women with incidental fibroids would not have multiple visits with ICD-9 codes for fibroids. Also, we modified the criteria to include women who had visits with an ICD-9 diagnosis of fibroids during the specified study period (April 4, 2014 to October 1, 2015), but whose imaging reports showing uterine fibroids had occurred up to one year prior to the study period. Our premise was that women who were symptomatic would have documentation confirming the presence of fibroids over a longer period of time. Algorithm 2 identified 1,174 potential cases of symptomatic uterine fibroids. Manual review of a subset of 51 EHRs found that 50 (98%) of the 51 records had image reports that confirmed the presence of uterine fibroids. Manual chart review yielded 30 true cases of symptomatic uterine fibroids, with a PPV of 65% (95% Confidence Interval: 0.5–0.79). Some cases were misclassified due to fibroids identified during emergency department visits for non-gynecologic conditions (e.g. trauma) or among pregnant women during routine perinatal ultrasound. These women truly had fibroids but they were not deemed as “symptomatic” upon expert manual review.

In Algorithm 3, we excluded women with incidental fibroids diagnosed during pregnancy. Women who had a positive pregnancy test (laboratory data) within three months of the fibroid diagnosis or fibroids detected through prenatal rather than general ultrasound were excluded. In an effort to exclude women with incidental fibroids diagnosed at the time of an emergency department (ED) visit or other ambulatory visit for a non-gynecologic symptom, we required that women have at least two ICD-9 diagnosis codes for fibroids on separate visit (non ED visit) dates. For Algorithm 3, 99% of the charts reviewed contained image documentation of uterine fibroids. After review of 300 EHRs, Algorithm 3 had a PPV of 76% (95% CI: 71–81).

## DISCUSSION

To the best of our knowledge, this is one of the first efforts to develop a computable phenotype to identify women with symptomatic uterine fibroids. Using an iterative process, we refined the algorithms to exclude asymptomatic cases, including incidental fibroids diagnosed during pregnancy, at the time of clinical evaluation of other gynecologic conditions or during non-gynecologic related emergency department evaluations. The final computable phenotype had a PPV of 76%. This performance seems acceptable given that there is no standard clinical variables or laboratory values for identifying symptomatic uterine fibroids. Relying solely on an automated algorithm for symptomatic fibroids may not be possible. However, our purpose was not to use the algorithms exclusively, but rather, as a first step in identifying women with symptomatic uterine fibroids for screening and potential enrollment.

While EHR-derived algorithms to simply identify cases of uterine fibroids have been associated with a sensitivity, specificity and PPV

of over 95%,[12] much less is known about the ability of EHR algorithms to discriminate between patients who are symptomatic and patients with asymptomatic, incidentally found fibroids. Our final algorithm has a PPV of 99% for image-confirmed fibroids (Table 2) that is similar to that reported by Feingold-Link and colleagues[12] to identify patients with image-confirmed fibroids in a health system database.

The results of this work can inform the use of EHR data for use in conducting large-scale fibroid research studies. From a clinical perspective, the final algorithm is simple and able to be easily implemented within a healthcare system for the conduct of large research studies to assess long-term health or treatment outcomes. The final computable phenotyping algorithm can serve as a practical first step in identifying women who are likely to be symptomatic and qualify for recruitment to research registries and intervention studies. From a population health perspective, the algorithm can contribute to other on-going methods to identify symptomatic patients to assess variations in surgical treatment procedures in patients across demographic, socio-economic, and geographical strata.[16,17]

The prevalence of symptomatic fibroids could have affected the PPV of the final algorithm. Patient self-report of fibroid symptoms and the completeness of documentation[18] of symptoms by the clinician are two factors that may also have affected the PPV. It is possible that patients with milder symptoms may not have discussed them with their provider. Clinicians may not have clearly documented symptoms if it was not the primary reason for the health care visit or alternatively, if symptoms were part of several complaints or inaccurately attributed to other co-morbidity.

Strengths of this study include our ability to use a combination of billing and EHR data from a large health care system. Data was obtained from a well-established data source warehouse in which prior algorithms have been developed to identify a range of disease conditions for observational studies and registry recruitment.[19] The components of the algorithm are simple and able to be applied to other health care systems.

There are several limitations that deserve attention. The analysis is based, in part, on administrative billing data that may have coding errors. However, fibroids are a common gynecologic problem with diagnosis codes that are used regularly by clinicians. Review of a subset of EHRs found that the majority of women with an ICD-9 diagnosis code for fibroids had image-confirmed fibroids at chart review. Our study was initiated prior to the implementation of ICD-10 codes at our institution. We therefore, maintained a timeframe (April 4, 2014 to October 1, 2015) that would maintain consistency in using data which was coded using ICD-9 codes. We developed the algorithms within a single academic health center which may limit the generalizability of the findings to other academic or community health systems. Coding practices within UNCHCS may differ from other academic or community health systems. We plan further testing of the algorithm in our own health system at different time points and in other health organizations to inform refinements for broader use in fibroid research. Finally, we were unable to evaluate algorithm sensitivity and

specificity because resource constraints limited our ability to randomly select a new set of charts (~300) from the EHR to reapply our algorithm.

## CONCLUSIONS

We developed and validated a computable phenotype with a PPV of 76% that can begin to identify women with symptomatic uterine fibroids using a combination of clinical and billing data and expert chart review. This work represents a first step to identify, via an EHR computable phenotype, patients with a highly prevalent condition that is associated with substantial morbidity and lower perceived quality of life. Further this approach may help to enhance research on the effectiveness of management options, treatment patterns and long-term outcomes. If EHRs can be used effectively to identify patients with symptomatic disease for enrollment into a research registry to evaluate the clinical and quality of life outcomes of a surgical or medical treatment, then clinicians and researchers will have a powerful tool for designing surveillance systems of population health and patient-centered outcomes.

## Acknowledgments

This project was supported by grant number P50HS023418 from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the authors and does not necessarily represent the official views of from the Agency for Healthcare Research and Quality.

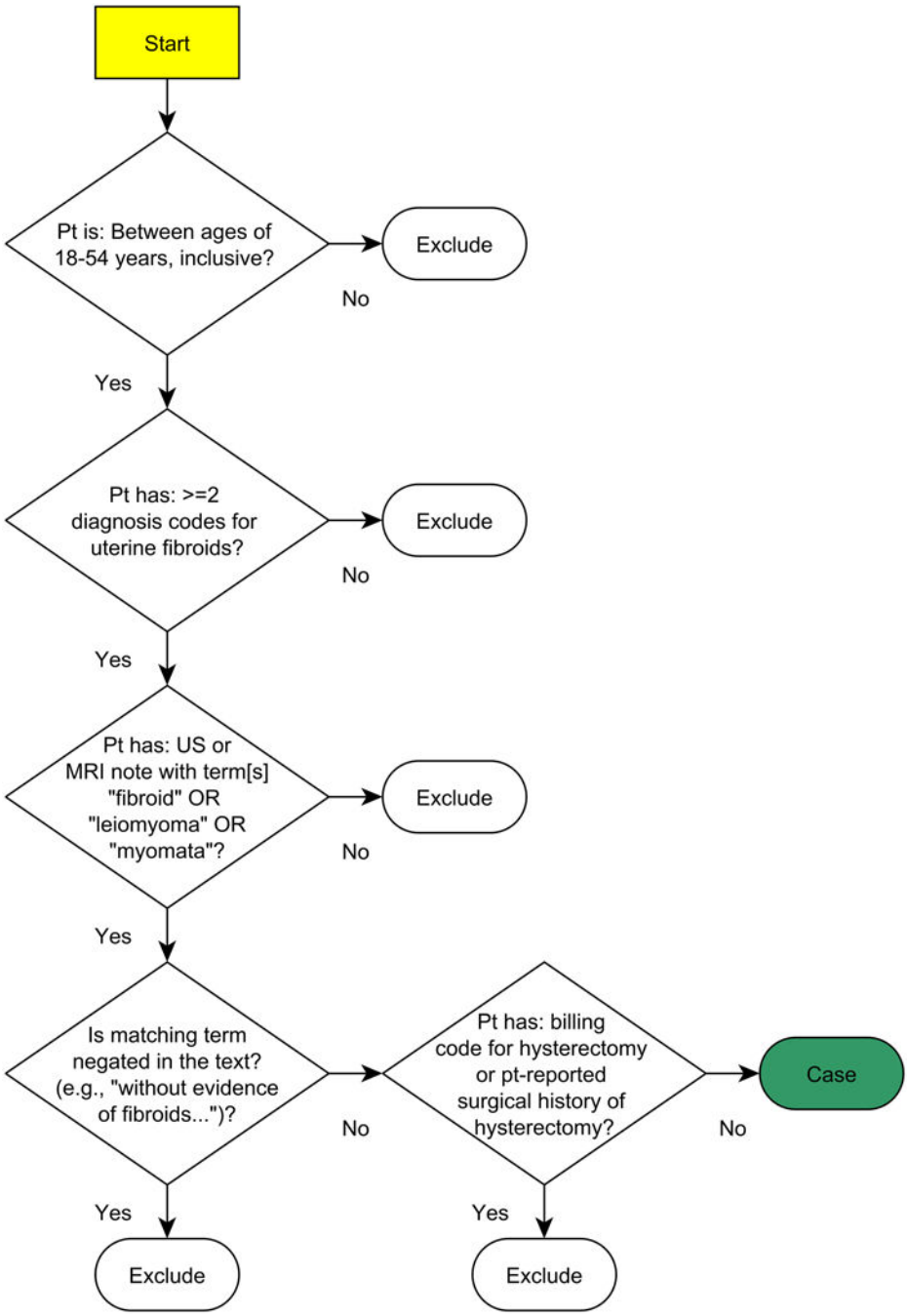
A portion of the findings were presented at the 32nd International Conference on Pharmacoepidemiology & Therapeutic Risk Management, Dublin, Ireland, August 25–28, 2016

## REFERENCES

1. http://sites.duke.edu/rethinkingclinicaltrials/informed-consent-in-pragmatic-clinical-trials/. Accessed July 9,2017 NHCSRCEHR-BPIRCTALToPCTAa.
2. Jha AK. Meaningful use of electronic health records: the road ahead. Jama. 10 20 2010;304(15): 1709–1710. [PubMed: 20959581]
3. Richesson RL, Hammond WE, Nahm M, et al. Electronic health records based phenotyping in next-generation clinical trials: a perspective from the NIH Health Care Systems Collaboratory. Journal of the American Medical Informatics Association: JAMIA. 12 2013;20(e2):e226–231. [PubMed: 23956018]
4. Newton KM, Peissig PL, Kho AN, et al. Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network. Journal of the American Medical Informatics Association: JAMIA. 6 2013;20(e1):e147–154. [PubMed: 23531748]
5. Khan AT, Shehmar M, Gupta JK. Uterine fibroids: current perspectives. International journal of women's health. 2014;6:95–114.
6. Laughlin SK, Schroeder JC, Baird DD. New directions in the epidemiology of uterine fibroids. Seminars in reproductive medicine. 5 2010;28(3):204–217. [PubMed: 20414843]
7. Borah BJ, Nicholson WK, Bradley L, Stewart EA. The impact of uterine leiomyomas: a national survey of affected women. American journal of obstetrics and gynecology. 10 2013;209(4):319 e311–319 e320. [PubMed: 23891629]
8. Stewart EA, Nicholson WK, Bradley L, Borah BJ. The burden of uterine fibroids for African-American women: results of a national survey. Journal of women's health. 10 2013;22(10):807–816.
9. Gliklich RE LM, Velentgas P, Campion DM, Mohr P, Sabharwal R, Whicher D, Myers ER, Nicholson WK. Identification of Future Research Needs in the Comparative Management of Uterine Fibroid Disease. A Report on the Priority-Setting Process, Preliminary Data Analysis, and Research Plan Effective Healthcare Research Report No. 31. (Prepared by the Outcome DEcIDE Center,

under Contract No. HHSA 290–2005–0035-I, TO5). AHRQ Publication No. 11- EHC023-EF. Rockville, MD: Agency for Healthcare Research and Quality 3 2011 Available at: http://effectivehealthcare.ahrq.gov/reports/final.cfm.

10. http://www.pcori.org/research-results/2014/comparing-options-management-patient-centered-results-uterine-fibroids-compare. Published 2016 P-CORICOfMP-CRfUFC-U.
11. http://compare-uf.org/ (Last accessed on July 23.
12. Feingold-Link L, Edwards TL, Jones S, Hartmann KE, Velez Edwards DR. Enhancing uterine fibroid research through utilization of biorepositories linked to electronic medical record data. Journal of women's health. 12 2014;23(12):1027–1032.
13. https://tracs.unc.edu/index.php/services/biomedical-informatics/cdw-h CDWfHC-HAa.
14. Levens ED, Wesley R, Premkumar A, Blocker W, Nieman LK. Magnetic resonance imaging and transvaginal ultrasound for determining fibroid burden: implications for research and clinical care. American journal of obstetrics and gynecology. 5 2009;200(5):537 e531–537. [PubMed: 19268886]
15. Abraham J SK. Computation of CIs for Binomial proportions in SAS and its practical difficulties. PhUSE. 2013:SP05:01–08.
16. Borah BJ, Laughlin-Tommaso SK, Myers ER, Yao X, Stewart EA. Association Between Patient Characteristics and Treatment Procedure Among Patients With Uterine Leiomyomas. Obstetrics and gynecology. 1 2016;127(1):67–77. [PubMed: 26646122]
17. Laughlin-Tommaso SK, Jacoby VL, Myers ER. Disparities in Fibroid Incidence, Prognosis, and Management. Obstetrics and gynecology clinics of North America. 3 2017;44(1):81–94. [PubMed: 28160895]
18. Jones SS, Adams JL, Schneider EC, Ringel JS, McGlynn EA. Electronic health record adoption and quality improvement in US hospitals. The American journal of managed care. 12 2010;16(12 Suppl HIT):SP64–71. [PubMed: 21314225]
19. Zhong VW, Obeid JS, Craig JB, et al. An efficient approach for surveillance of childhood diabetes by type derived from electronic health record data: the SEARCH for Diabetes in Youth Study. Journal of the American Medical Informatics Association: JAMIA. 11 2016;23(6):1060–1067. [PubMed: 27107449]



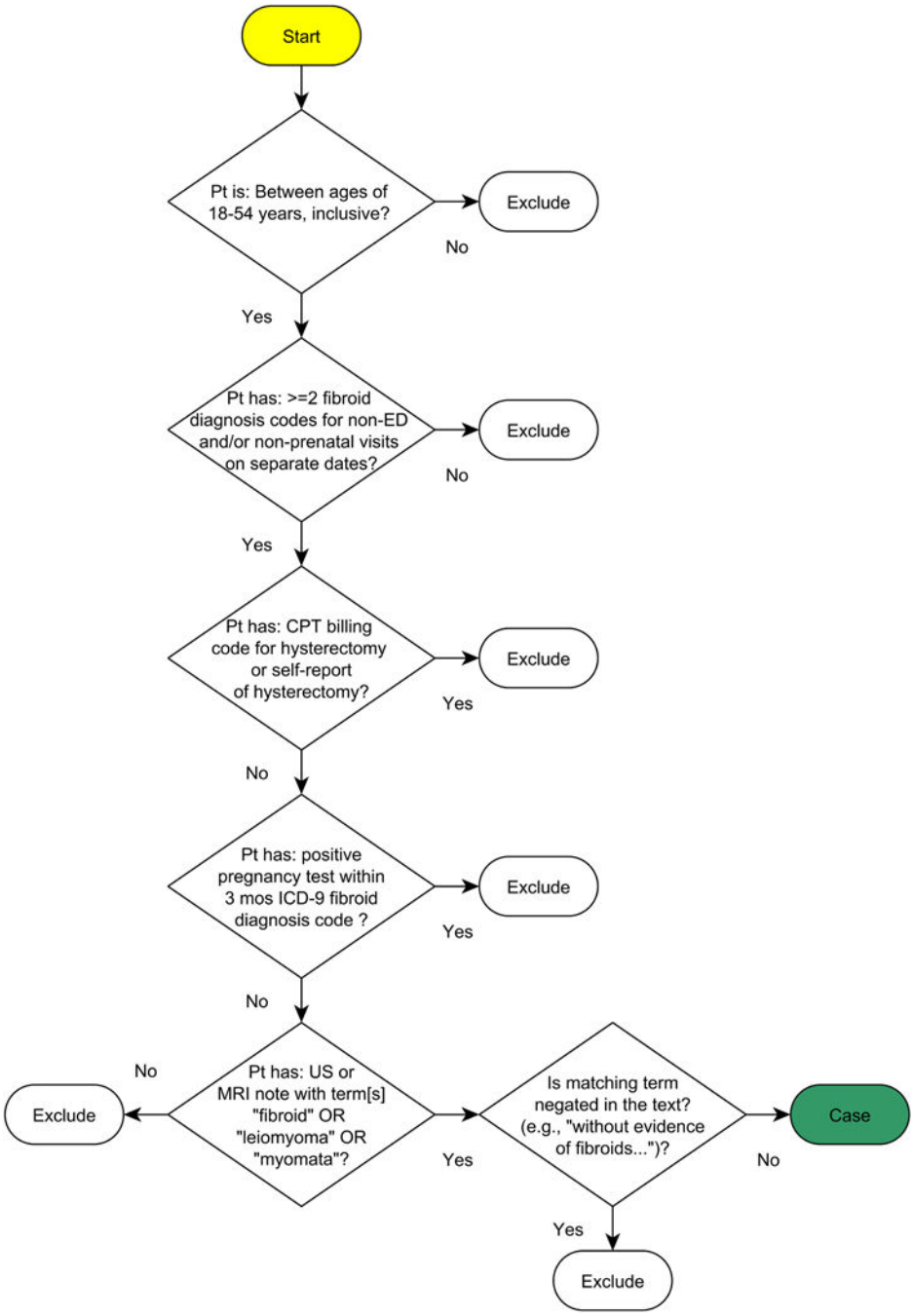



**Figure 1.**

Computable phenotyping algorithms to identify women with symptomatic uterine fibroids in an electronic health record (EHR) database

Algorithm 1. Initial algorithm for the identification of women with symptomatic uterine fibroids. CPT= current procedural terminology; Text-mining was used to identify keywords (fibroids, leiomyoma) in the imaging and clinical reports. Pt= patient; MRI = magnetic resonance imaging; US- pelvic ultrasound.

Algorithm 2. CPT= current procedural terminology. MRI: magnetic resonance imaging; OB: obstetrical; Pt: patient; US: ultrasound.

Algorithm 3. Outpatient visits included outpatient scheduled office visits. ED=emergency department; MRI: magnetic resonance imaging; OB: obstetrical; Pt: patient; US: ultrasound.

**Table 1.**

List of ICD-9 diagnosis and procedure codes and CPT codes used in the algorithms

| Description | Code Type | Codes | |
|---|---|---|---|
| Uterine fibroids (also known as fibroids, leiomyomas, myomas) | ICD-9 | 218.0 | |
| | | 218.1 | |
| | | 218.2 | |
| | | 218.9 | |
| Hysterectomy | ICD-9 | 68.3 | 68.6 |
| | | 68.31 | 68.61 |
| | | 68.39 | 68.69 |
| | | 68.4 | 68.7 |
| | | 68.41 | 68.71 |
| | | 68.49 | 68.79 |
| | | 68.5 | 68.8 |
| | | 68.51 | 68.9 |
| | | 68.59 | |
| Hysterectomy | CPT | 58150 | 58541 |
| | | 58152 | 58542 |
| | | 58180 | 58543 |
| | | 58200 | 58544 |
| | | 58210 | 58548 |
| | | 58240 | 58550 |
| | | 58260 | 58552 |
| | | 58262 | 58553 |
| | | 58263 | 58554 |
| | | 58267 | 58570 |
| | | 58270 | 58571 |
| | | 58275 | 58572 |
| | | 58280 | 58573 |
| | | 58285 | 58951 |
| | | 58290 | 58954 |
| | | 58291 | 58956 |
| | | 58292 | 59135 |
| | | 58293 | 59525 |
| | | 58294 | |

ICD- 9 = International Classification of Diseases, Ninth Revision; CPT= Current Procedural Terminology

**Table 2.**

Results of three computable phenotyping algorithms to detect symptomatic uterine fibroids using EHR and billing data from the UNCHS data warehouse, April 4, 2014-October 1, 2015

| Algorithms | No. presumptive cases (n) | Subset of charts reviewed (n) | Image confirmed fibroids in subset of charts, n (%) | Results from expert chart review | |
|---|---|---|---|---|---|
| | | | | Confirmed symptomatic fibroids[1] n (%) | Positive Predictive Value[2] (95% CI) |
| Algorithm 1 | 4,342 | 150 | 139 (93) | 71 (47) | 0.47 (0.39–0.56) |
| Algorithm 2 | 1,174 | 51 | 50 (98) | 30 (65) | 0.65 (0.50–0.79) |
| Algorithm 3 | 465 | 300 | 297 (99) | 225 (76) | 0.76 (0.71–0.81) |

Abbreviations: EHR = Electronic health record; UNCHS = University of North Carolina Health System; PPV= Positive predictive value.

[1] Confirmed with expert chart review.

[2] Positive predictive value was calculated based on the equation: Number of patients with symptomatic uterine fibroids confirmed by expert chart review divided by the number of patients identified by the algorithm as having symptomatic uterine fibroids. Chart reviews and PPV calculations were performed for a random sample of charts for each algorithm. For algorithms 2 and 3, there were five charts that contained insufficient information to determine if the confirmed fibroids were symptomatic. These charts were not counted in the denominator when calculating the PPVs for symptomatic fibroids for each algorithm.