



HHS Public Access

Author manuscript

Am J Obstet Gynecol. Author manuscript; available in PMC 2023 March 01.

Published in final edited form as:

Am J Obstet Gynecol. 2022 March ; 226(3): 388.e1–388.e11. doi:10.1016/j.ajog.2021.11.020.

Development of an algorithm to assess unmeasured symptom severity in gynecologic care

Kemi M. Doll, MD, MSCR,

Department of Obstetrics & Gynecology, University of Washington School of Medicine, Seattle, WA, USA; Department of Health Services, University of Washington School of Public Health, 1959 NE Pacific St. Box 356460, Seattle, WA 98195

Annie Green Howard, PhD,

Department of Biostatistics, Gillings School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA., 123 West Franklin St. CB #8120, Chapel Hill, NC 27516

Till Stürmer, MD, MPH, PhD,

Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, 2105 McGavran-Greenberg Hall, CB #7435, Chapel Hill, NC 27599

Tim Carey, MD, MPH,

Department of Medicine, UNC School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; The Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, 5034 Old Clinic Building, CB# 7110, Chapel Hill, NC 27599

Corresponding Author: Kemi M. Doll, Kdoll@uw.edu, 1959 NE Pacific St. Box 356460, Seattle, WA 98195.

Kemi M. Doll: Conceptualization, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing, Visualization, Supervision, Funding Acquisition

Annie Green Howard: Methodology, Formal Analysis, Data Curation, Writing – Review & Editing, Visualization, Funding Acquisition

Till Stürmer: Writing – Review & Editing

Tim Carey: Writing – Review & Editing

Wanda K. Nicholson: Writing – Review & Editing **Erin Carey:** Investigation, Writing – Review & Editing **Evan Myers:** Writing – Review & Editing

David Nerenz: Writing – Review & Editing

Whitney R. Robinson: Conceptualization, Methodology, Investigation, Writing – Review & Editing, Supervision, Funding Acquisition

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The authors report no conflict of interest.

Research reported in this publication was supported by *the National Institute of Minority Health and Health Disparities* of the National Institutes of Health under award number 1R01MD011680. The funder had no role in the study design, collection, analysis or interpretation of data, in the writing of the report, or in the decision to submit the article for publication.

List of Supplemental Digital Content Files

1. ICD9, ICD 10, and CPT Codes from Dataset Construction
2. Data Analysis and Creation of Scoring Rubric Example: Vaginal Bleeding
3. Distributions of Symptom Presence/Absence By Composite Bulk Symptom Severity Score
4. Distributions of Symptom Presence/Absence By Composite Bleeding Symptom Severity Score
5. Distributions of Symptom Presence/Absence By Composite Pain Symptom Severity Score

Wanda K. Nicholson, MD, MPH, MBA,

Center for Women's Health Research, Department of Obstetrics and Gynecology, UNC School of Medicine, Chapel Hill, NC, USA; Center for Health Promotion and Disease Prevention, UNC School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA., 3027 Old Clinic Building, Campus Box #7570, Chapel Hill, NC, 27599

Erin Carey, MD, MSCR,

Department of Obstetrics and Gynecology, University of North Carolina School of Medicine, Chapel Hill, NC, USA., 4010 Old Clinic Building, CB# 7570, Chapel Hill, NC 27599

Evan Myers, MD,

Division of Reproductive Sciences, Department of Obstetrics & Gynecology, Duke University School of Medicine, Durham, NC, USA., 244 Baker House, Durham, NC 27710

David Nerenz, PhD,

Center for Health Policy and Health Services Research, Henry Ford Health System, Detroit, MI, USA., Ste 3 A., One Ford Place, Detroit, MI 48202

Whitney R. Robinson, PhD, MSPH

Department of Epidemiology, UNC Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; 2104-B McGavran-Greenberg Hall, CB #7435, Chapel Hill, NC 27599; Division of Women's Community and Population Health, Department of Obstetrics and Gynecology, Duke University School of Medicine, 200 Trent Drive, 203 Baker House, DUMC 3084, Durham, NC 27710

Abstract

Background: Disparities research is often limited by incomplete accounting for differences in health status by populations. In the U.S., hysterectomy shows marked variation by race and geography, but it is difficult to understand what factors cause these variations without accounting for differences in the severity of gynecologic symptoms that drive hysterectomy decision-making.

Objective: Our objective is to demonstrate a method for using electronic health record (EHR)-derived data to create composite symptom severity indices to more fully capture relevant markers that influence the decision for hysterectomy.

Study Design: This was a retrospective cohort study of 1,993 people who underwent hysterectomy between April 4, 2014 and December 31, 2017 from ten hospitals and over 100 outpatient clinics in North Carolina. EHR data including billing, pharmacy, laboratory data, and free text notes, were used to identify markers of 3 common indications for hysterectomy: bulk symptoms (pressure from uterine enlargement), vaginal bleeding, and pelvic pain. To develop weighted symptom indices, we finalized a scoring algorithm based on the relationship of each marker to an objective measure, in combination with clinical expertise, with the goal of composite symptom severity indices that had sufficient variation to be useful in comparing different patient groups and allow discrimination between more or less severe symptoms of bulk, bleeding, or pain.

Results: Ranges of symptom severity scores varied across the three indices; including composite bulk score (0 to 14), vaginal bleeding score (0 to 44) and pain score (0 to 30). Mean values of each

composite symptom severity index were greater for those who had diagnostic codes for vaginal bleeding, bulk symptoms, or pelvic pain, respectively. However, each index also demonstrated variation across the entire group of hysterectomy cases and identified symptoms that ranged in severity among those with and without the target diagnostic codes.

Conclusions: Leveraging multisource data to create composite symptom severity indices provided greater discriminatory power to assess common gynecologic indications for hysterectomy. These methods can advance understanding in healthcare use in the setting of long-standing inequities and be applied across populations to account for previously unexplained variation across race, geography, and other social indicators.

Keywords

Hysterectomy; Health Equity; Electronic Health Records; Leiomyoma; Quality of Life

Introduction

Efforts to measure patient-reported symptoms are essential to better patient care and to identifying the drivers of unexplained racial differences in symptomatology and treatment. Important medical indications for treatment may either not be well-measured or such measurements can be biased by race or other social factors because of how they were originally constructed.^{1,2}

Hysterectomy for benign disease is both common and marked with inequalities by race, ethnicity, insurance status, and geography.^{3–11} For the common symptoms of uterine bulk, vaginal bleeding, and pelvic pain, there are other treatments available, and it is the severity of the symptom alongside failure of prior medical treatments that drive appropriateness of hysterectomy. Within electronic health records (EHR), however, there is a lack of standardization of assessing symptom severity and no commonly used patient reported outcome measures for these symptoms.

In a context in which decision-making is informed by patient-reported symptoms, quality of life effects, and shared decision-making tools,^{12–15} diagnostic codes and lab values alone are insufficient to account for differences in patient symptom severity and have failed to explain marked racial variation.¹⁶ In addition, the cumulative effects of symptoms over time or treatments previously tried and failed go unmeasured. Therefore, although numerous studies have documented variation in hysterectomy use, it has remained unclear to what extent such variation represents differences in clinical indication, symptom severity, patient preferences, or biases within the health care system.

In this paper we demonstrate an approach to characterize gynecologic symptom severity in a racially and socioeconomically diverse sample of over 1,900 premenopausal people treated with hysterectomy. Our objective was to use data from administrative billing and the EHR from 10 hospitals and follow a rigorous, multi-step process to construct symptom severity indices for the three most prominent gynecologic indications for hysterectomy: bulk (pressure) symptoms from uterine enlargement, vaginal bleeding, and pelvic pain.^{9, 17}

Materials and Methods

Data Sources.

The Carolina Data Warehouse for Health (CDW-H) is a searchable federation of electronic health information and administrative data from the University of North Carolina (UNC) Health system, with information from the 10 hospitals and hundreds of affiliated practices. We queried the CDW-H for structured clinical data and supplemented this structured data with free text and images from the EHR, captured by a team of professional abstractors.

Cohort identification using CDW-H.

As part of a larger study to examine determinants of racial disparities in pre-menopausal hysterectomy, those eligible for the cohort included North Carolina residents age 18 to 44 years old who underwent hysterectomy for benign (non-cancer related) disease between April 4, 2014 and December 31, 2017. The upper age limit of 44 was chosen as a conservative cut-point to identify premenopausal people with a high degree of specificity, as menopausal status is a major determinant of surgical decision-making and <5% of people with female reproductive organs undergo natural menopause before age 45 years.¹⁸ Administrative billing codes were used to identify all hysterectomies performed (See Supplemental Digital Content 1) during the eligible date range, including an ICD9 to ICD10 crosswalk to ensure full capture. Included sites had to have implemented Epic at least 180 days before the patient's surgery. People were excluded if they were pregnant at the time of surgery, were not a North Carolina resident, had prior or active breast, ovarian, uterine, or cervical cancer diagnoses, or cancers with treatment plans that may involve hysterectomy (bladder, anal, colorectal). The average follow-back time for the analysis sample was 691 days (standard deviation of 345).

Capture of Structured Data: Carolina Data Warehouse for Health.

Sociodemographic, clinical, and laboratory data were collected from the CDW-H. For sociodemographic factors, we captured date of birth and a 6-level race variable (White, African American/Black, Asian, American Indian/Alaska Native, Other, Refused/Unknown), Hispanic/Latino ethnicity (yes, no) from Epic, age, height and weight, marital status, home address, and insurance at the time of hysterectomy. For clinical information, we captured date of surgery, all physician-billed and hospital-billed diagnostic and procedure codes associated with the hysterectomy encounter, all hemoglobin (Hgb) values, blood transfusions, imaging procedures, gynecologic well-care visits, prescriptions for pain medication, Emergency Department visits, and hospital admissions up to 12 months prior to surgery, and all diagnostic codes at the time of hysterectomy for the primary symptoms of interest: vaginal bleeding, pelvic pain, and bulk symptoms. There were several codes that mapped to each symptom to capture all potential symptom report (See Supplemental Digital Content 1).

Capture of Unstructured Data: EHR Abstraction.

EHR Abstraction: Overview and Rationale.—We created an EHR data abstraction tool in RedCAP and accompanying protocol to capture candidate markers of symptom

severity: the presence of gynecologic diagnoses, symptom descriptions, and surgeon-reported indication for hysterectomy for up to 12 months prior to surgery. Candidate markers of symptom severity were based on prior literature^{19–23} and expert clinical input from the study team (K.D., E.C., E.M., W.N.). In addition to the presence or absence of specific sequelae of vaginal bleeding, pelvic pain, and abdominal/pelvic bulk symptoms, we captured health care utilization data (ER visits, blood transfusions, opioid and other pain medication use), as well as missed days of work/activity. We planned overlap with several data points also captured by the CDW-H structured data to capture possible events and services completed outside of the UNC system but documented in healthcare provider notes.

EHR abstraction: Pilot study to test and refine abstraction tool.—Before finalizing the EHR data abstraction protocol, we conducted a pilot study to assess and refine it. We evenly sampled 52 cases among the hospital sites (random sample of 5–6 records per site), and among the three symptoms of interest, identified by relevant diagnostic codes, and followed the abstraction protocol. Based on this experience, the protocol was then updated to eliminate needlessly repetitive information with the structured data (e.g., Lab values for Hgb and opioid prescriptions). We also found planned data elements with missing rates too high to be of meaningful use (e.g., tampon/pad count, documented in only 4 cases; inability to tolerate an exam, documented in 0 cases) and eliminated them for abstraction efficiency. We did not collect 0–10 pain scores as they were not consistently captured in notes nor in structured data, and not specific to any organ site or group. We had three possibilities for symptom report – present, absent, or absent from the record. In the pilot, only one symptom in one case was marked as ‘absent,’ and the rest were either present or not commented upon. Therefore, we adjusted data entry to be “Yes” or “Absent from record.” The protocol was then finalized and full abstraction completed.

EHR abstraction: Quality Control.—During the abstraction, we adhered to the protocol guide and kept a corollary log; together, these served as active documents with auditable updates based on abstractors’ feedback and ongoing quality assurance review. A team of four abstractors with over 20 years of cumulative experience completed all data abstraction. They could initiate secondary review for any data ambiguity, and 5% of all records from each site were chosen at random for double abstraction by an abstractor with clinical experience or the abstractor team lead to ensure accuracy.²⁴ Discrepancies flagged by abstractors (3% of data fields out of 100 unique records that were randomly sampled for quality check) were resolved through group discussion in consultation with clinical leads (K.D., E.C) followed by appropriate protocol updates. All data was abstracted into REDCap and merged with administrative data for the final analytic dataset.

EHR abstraction: Categorizing free text captured by abstractors.—While most fields in the REDCap abstraction tool forced data to be recorded in a structured format (i.e., numerical, yes/absent), there were some fields that allowed abstractors to record free text when they were unsure if the default options applied. All free text entries were reviewed and either re-coded into existing abstraction categories (e.g., ‘pelvic floor tension myalgia’ recoded into existing ‘pelvic pain’ diagnosis) or used to create new categories not previously

identified. Abstracted free-text data were also used to re-apply exclusion criteria for cancer and pregnancy that administrative code definitions missed.

Construction of weighted indices for symptom severity.

Our goal was to create a composite index for each symptom – bleeding, bulk, and pain – that was comprised of appropriate symptom markers, weighted by their relative severity. First, we created histograms and descriptive tables for all symptom severity candidate markers - gynecologic diagnoses, symptom descriptions, and surgeon-reported indication for hysterectomy. The weights for each marker were first assigned by the investigators' clinical expertise with higher weights for markers more severe and more rare. For example, the presence of a report of 'heavy bleeding' was given 1 point, iron supplementation was given 3 points, and history of blood transfusion for non-surgery related anemia was given 5 points.

Second, each marker was compared against an objective measure of severity. The objective measures were uterine weight from pathology report, the presence of anemia by lab criteria, and the presence of opioid prescriptions for bulk, vaginal bleeding, and pelvic pain, respectively. We examined associations by visual inspection of histograms and sample distribution. Markers that were more strongly associated with objective measures of symptom severity were more highly weighted to improve the construct validity of the final composite symptom severity indices. In addition, a few candidate symptom markers (urinary symptoms, constipation, and weight gain) were not meaningfully associated with any of the 3 objective measures, were considered non-specific, and so not included in the final three composite symptom severity indices.

Third, we then noted administrative coding patterns that were indicative of more severe values of the objective criteria and therefore gave them higher weights (points). For example, the presence of the ICD diagnostic code for uterine hypertrophy (which does not specify uterine size) in the year *prior* to surgery was uncommon and associated with larger uterine weight compared to such coding only present on the surgery encounter.

Fourth, we chose to include the objective measures in each respective symptom severity index, as these measures are also important markers of symptom severity. By excluding them from the final scoring system, we would be omitting key data, and undermeasuring severity in future use.

Refinement of weighted indices for symptom severity.

We completed several index iterations to optimize variability while maintaining logical progression of scoring from least to most severely symptomatic. Supplemental Digital Content 2 demonstrates the step-by-step process for the Vaginal Bleeding score, as an example. We tested whether the average followback time varied according to severity score to assess whether those with higher symptom severity scores were reflecting more follow-back time in which an individual's symptoms could be captured. Look-back time available did not vary by bleeding severity score ($p=0.61$), or pain severity score ($p=0.86$). It had a tendency to be *smaller* for those with higher bulk scores (717 for $\geq 75^{\text{th}}$ percentile) than for those with low bulk scores (860 for $< 25^{\text{th}}$ %ile). This is the opposite of the pattern one would expect if shorter follow-up time was biasing our estimates by missing

symptom markers. At no point during this process were data stratified or analyzed by any demographic factor, including race.

Data Analysis: Evaluation of weighted indices for symptom severity compared to symptom specific diagnostic codes.

Descriptive statistics were used to assess agreement between each symptom severity index score with the diagnosis code of that symptom at time of surgery – as evidenced on hospital billing for the procedure. Due to the skewed distribution of the index scores, we tested for a difference in severity score between diagnosis present or absent using the Wilcoxon Rank Sum Test.

This study received approval from the University of North Carolina's Institutional Review Board on 11/06/2017 (study ID: 17–2728).

Results

An initial 2,830 individuals were identified through the CDW-H that were aged 18 to 44 years and had a hysterectomy within the study timeframe. Of these, 1,933 met inclusion criteria by query of the structured data and underwent EHR abstraction. With removal of duplicates, empty records, and those whose abstracted information met exclusion criteria, our final merged analytic cohort was 1,913 (Figure 1). About a quarter (425) of the 1,913 patients had a look-back period <365 days. The 1st percentile was 186 days, while the 10% percentile was 252 days.

Average follow-back time did not differ by age at time of surgery ($p=0.30$) or race/ethnicity ($p=0.15$). Ten individuals were missing uterine weight and 153 individuals were missing Hgb lab results, and were excluded from those respective symptom indices. Cohort characteristics and overall symptom prevalence are reported in Table 1.

Comparison of individual markers of symptom severity with symptom specific diagnostic codes.

Several markers of symptom severity did not fully overlap with the presence of the relevant symptom specific diagnosis code (Table 2). In some cases, people had reported symptoms in the EHR that were not captured in diagnostic codes. For example, 31% of women who did not have a diagnostic code for vaginal bleeding reported heavy bleeding as a problematic symptom from free text information in the EHR. Among those with the diagnostic code for vaginal bleeding, the presence of severe symptom markers (e.g., lightheadedness/dizziness (26%), or requiring blood transfusion (7%)) were present among a significant minority who would be indistinguishable from those with milder symptoms based solely on codes (Table 2).

Composite symptom severity indices.

The final composite symptom severity indices for vaginal bleeding, bulk symptoms, and pelvic pain are seen in Table 3, with the distribution of scores across the study population depicted in Figure 2, along with an overlay of the categorization when using diagnostic

codes alone. There were different ranges of scores across the three metrics with the composite bulk symptom severity score ranging from 0 to 14, composite vaginal bleeding symptom severity score ranging from 0 to 44 and composite pain symptom severity score ranging from 0 to 30.

Bulk index scores: comparison with bulk diagnostic codes

We found evidence (p -value < 0.0001) of differences in composite bulk symptom severity score between those with bulk diagnostic code present at surgery [median of 6 [IQR: (3,7)]] as compared to those without the code [median of 0 [IQR: (0, 1)]]. However, 9% of the 251 individuals with a composite bulk symptom severity index score at or above 6 did not have a bulk diagnosis code reported at surgery. Additionally, as an example of the heterogeneity in symptoms, there were a total of 6 different combinations of symptoms that resulted in a composite bulk symptom severity index score of 7. One person with a composite bulk symptom severity score of 7 reported bloating, had a uterine size between the 50th and 75th percentile, and a bulk diagnosis code in the year prior to surgery. Another person had the same score but had bulk NOS and a uterine size greater than or equal to the 75th percentile (Supplemental Digital Content 3 shows the distributions of symptom presence/absence by composite bulk symptom severity score).

Bleeding index scores: comparison with anemia diagnostic codes

Similarly, composite bleeding symptom severity index scores differed based on diagnostic code of vaginal bleeding at surgery: median composite bleeding symptom severity score was 7 [IQR: (4,12)] with the diagnostic code present and 1 [IQR: (0,4)] (p -value < 0.0001) with the diagnostic code absent. As with bulk, there are individuals with high composite bleeding symptom severity index scores who did not have a diagnosis code of bleeding at the time of surgery. Specifically, 13% of the 819 individuals with composite bleeding symptom severity scores at or above 7 did not have a bleeding diagnosis code recorded at surgery. We also observed variations in symptoms across composite bleeding symptom severity scores (Supplemental Digital Content 4).

Pain index scores: Comparison with pelvic pain diagnostic codes

Likewise, our composite pain symptom severity score differed by pain diagnosis at or in the year prior to surgery ($p < 0.0001$) with median composite pain symptom severity score among those with pain diagnosis present at surgery being 10 [IQR: (6,16)] compared to 3 [IQR: (0,7)] in those without the diagnosis. Of the 384 individuals, 59% with composite pain symptom severity scores at or above 10 did not have a pain diagnosis code at the time of surgery. We found evidence, as with the other severity scores, of variation in symptoms among individuals with the same composite pain symptom severity score (Supplemental Digital Content 5).

Comment

Principal Findings

We developed three composite symptom severity indices for the three most common indications for hysterectomy in the U.S. High scores on the indices were strongly associated

with presence of diagnostic codes. However, the indices also detected a substantial proportion of cases (9% to 59%) who were not coded for that symptom via diagnostic codes. Further, the ranges of the indices allow greater statistical discrimination among those with diagnosis codes present (ranging from 16% - 67%, Table 1). This is a needed step on the pathway to equity in hysterectomy in the U.S.

Clinical Implications

Without accounting for racial/ethnic differences in symptom severity, it is impossible to define a reasonable level of racial/ethnic difference in hysterectomy rates and design clinical interventions to achieve that target. We interpret the discordance between diagnostic codes and symptom severity as arising from the fact that billing codes may just capture what is necessary for payment approval as distinct to the full experience of the patient, an important limitation when they are alone used to define gynecologic symptom status of an individual. Without a greater ability to account for currently undermeasured aspects of symptom severity in gynecology, our assessments of care quality of hysterectomy and other procedures do not take into account a major component of clinical decision making. With the development of these measures, we have a path forward to evaluate, update, revise and/or restructure current clinical guidance to ensure equitable distribution of hysterectomy among premenopausal people.

Research Implications

This work is a part of a larger multi-year project to investigate causes of hysterectomy rate disparity in the U.S. South. Given the extensive processes required to develop and refine these indices, this method is presented separately here. In addition, we created these measures blind to any racial/ethnic categorization to minimize bias for the larger study, whose analysis is underway. With these indices we may move forward to uncover specific etiology and therefore appropriate intervention for longstanding hysterectomy disparity. We can examine how symptom severity may vary by groups of interest, indicating either a clinical need for increased access to and development of uterine-sparing treatments or a systemic retraining on how symptom severity may be differentially assessed and acted upon. These indices are therefore not limited to analyses of racial disparity, but can be utilized in any research that seeks to account for gynecologic symptom burden as an important influential factor in the outcome of interest. This includes other racial/ethnic populations, groups defined by language or nativity, SES, insurance status, sexual and gender minorities, and those with disabilities.

Results in the Context of What is Known

Our current ability to execute interventions to ensure equity in hysterectomy treatment is limited by the fact that existing research does not sufficiently account for patient symptom severity. Without this information, the causes of race/ethnicity-based differences in treatment have remained contested over decades, with little productive movement toward a clear consensus. As we have demonstrated in this analysis using multi-source data, the severity of gynecologic symptoms can vary dramatically for the same diagnosis, from non-existent to disabling. Our final algorithms demonstrate finer ability to differentiate among degrees

of severity than the common use of administrative billing diagnostic codes, lab values, or pathology records alone.

Strengths and Limitations

One limitation with this approach is that continued variation, in the form of differing symptom aspects, still exists within each level of symptom severity score, which may indicate the limit of transformation of patient experience into quantitative data. Some people with same symptom severity will have higher scores just because their providers document better. So instead we conceptualize the indices as measures that are specific but still have some limitations when it comes to sensitivity -- better than codes alone, but not perfect. In addition, our composite score ranges differ (from 14 to 40) across the different composite symptom severity indices. In future use, scaling to a uniform 1–100 range would avoid differential weights in predictive analyses.

The parent study cohort was defined by performance of premenopausal hysterectomy, a salient clinical procedure with well-documented but unexplained differences in rates of treatment. Given that, these indices are useful for other studies of hysterectomy cohorts, but would need to be modified for non-hysterectomy cohorts given the inclusion of administrative data at the time of surgery and post-operative findings of uterine size. We acknowledge that EHR abstraction is expensive in both time and resources, therefore our method may not be easily scalable. We hope that advancements in natural language processing can make this process more efficient and accessible. These data support the value and feasibility of developing gynecologic-specific structured reporting of these common symptoms, especially in relation to hysterectomy decision making, which is increasingly possible with the customization of current major EHR vendors.

Conclusions

When put into use in healthcare disparities research, our proposed severity indices will provide critical tools to account for differences in patient health status. The ability to account for variation in the severity of patient symptoms will advance work that seeks to identify what factors drive differences in health care utilization and outcomes. Patient-centered, health equity literature cannot progress until the field develops better methods for characterizing clinical indications of complex clinical phenomena.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

Research reported in this publication was supported by *the National Institute of Minority Health and Health Disparities* of the National Institutes of Health under award number 1R01MD011680. This research is also supported by the National Institute of Nursing Research of the National Institutes of Health under award number F31NR018786. We acknowledge the editorial assistance of the NC Translational and Clinical Sciences (NC TraCS) Institute, which is supported by the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health, through Grant Award Number UL1TR002489.

References

1. Ledford H Millions of black people affected by racial bias in health-care algorithms. *Nature* 10 2019;574(7780):608–609. doi:10.1038/d41586-019-03228-6 [PubMed: 31664201]
2. Madhusoodanan J Is a racially-biased algorithm delaying health care for one million Black people? *Nature Dec* 2020;588(7839):546–547. doi:10.1038/d41586-020-03419-6 [PubMed: 33328647]
3. Callegari LS, Katon JG, Gray KE, et al. Associations between Race/Ethnicity, Uterine Fibroids, and Minimally Invasive Hysterectomy in the VA Healthcare System. *Womens Health Issues* 2019 Jan - Feb 2019;29(1):48–55. doi:10.1016/j.whi.2018.08.005 [PubMed: 30293778]
4. LoGerfo JP. Variation in surgical rates: fact vs. fantasy. *N Engl J Med* Aug 1977;297(7):387–9. doi:10.1056/NEJM197708182970711 [PubMed: 406565]
5. Vayda E, Morison M, Anderson GD. Surgical rates in the Canadian provinces, 1968 to 1972. *Can J Surg* May 1976;19(3):235–42. [PubMed: 1277019]
6. Jacoby VL, Fujimoto VY, Giudice LC, Kuppermann M, Washington AE. Racial and ethnic disparities in benign gynecologic conditions and associated surgeries. *Am J Obstet Gynecol* Jun 2010;202(6):514–21. doi:10.1016/j.ajog.2010.02.039 [PubMed: 20430357]
7. Bower JK, Schreiner PJ, Sternfeld B, Lewis CE. Black-White differences in hysterectomy prevalence: the CARDIA study. *Am J Public Health* Feb 2009;99(2):300–7. doi:10.2105/AJPH.2008.133702 [PubMed: 19059854]
8. Powell LH, Meyer P, Weiss G, et al. Ethnic differences in past hysterectomy for benign conditions. *Womens Health Issues* 2005 Jul-Aug 2005;15(4):179–86. doi:10.1016/j.whi.2005.05.002 [PubMed: 16051109]
9. Weiss G, Noorhasan D, Schott LL, Powell L, Randolph JF, Johnston JM. Racial differences in women who have a hysterectomy for benign conditions. *Womens Health Issues* 2009 May-Jun 2009;19(3):202–10. doi:10.1016/j.whi.2009.03.001 [PubMed: 19447324]
10. Wright JD, Herzog TJ, Tsui J, et al. Nationwide trends in the performance of inpatient hysterectomy in the United States. *Obstet Gynecol* Aug 2013;122(2 Pt 1):233–41. doi:10.1097/AOG.0b013e318299a6cf [PubMed: 23969789]
11. Katon JG, Bossick AS, Doll KM, et al. Contributors to Racial Disparities in Minimally Invasive Hysterectomy in the US Department of Veterans Affairs. *Med Care* 12 2019;57(12):930–936. doi:10.1097/MLR.0000000000001200 [PubMed: 31730567]
12. The American College of Obstetricians and Gynecologists. ACOG Practice Bulletin Number 128: Diagnosis of Abnormal Uterine Bleeding in Reproductive-Aged Women 2012.
13. Zandstra D, Busser JAS, Aarts JWM, Nieboer TE. Interventions to support shared decision-making for women with heavy menstrual bleeding: A systematic review. *Eur J Obstet Gynecol Reprod Biol* Apr 2017;211:156–163. doi:10.1016/j.ejogrb.2017.02.026 [PubMed: 28273646]
14. The American College of Obstetricians and Gynecologists. Committee Opinion Number 557: Management of Acute Abnormal Uterine Bleeding in Nonpregnant Reproductive-Aged Women 2013.
15. The American College of Obstetricians and Gynecologists. Committee Opinion Number 701: Choosing the Route of Hysterectomy for Benign Disease 2017.
16. Doll KM, Dusetzina SB, Robinson W. Trends in Inpatient and Outpatient Hysterectomy and Oophorectomy Rates Among Commercially Insured Women in the United States, 2000–2014. *JAMA Surg* Sep 1 2016;151(9):876–7. doi:10.1001/jamasurg.2016.0804 [PubMed: 27168235]
17. Learman LA, Kuppermann M, Gates E, Gregorich SE, Lewis J, Washington AE. Predictors of hysterectomy in women with common pelvic problems: a uterine survival analysis. *J Am Coll Surg* Apr 2007;204(4):633–41. doi:10.1016/j.jamcollsurg.2007.01.006 [PubMed: 17382223]
18. Rosner B, Colditz GA. Age at menopause: imputing age at menopause for women with a hysterectomy with application to risk of postmenopausal breast cancer. *Ann Epidemiol* Jun 2011;21(6):450–60. doi:10.1016/j.annepidem.2011.02.010 [PubMed: 21441037]
19. Magnay JL, O'Brien S, Gerlinger C, Seitz C. A systematic review of methods to measure menstrual blood loss. *BMC Womens Health* 08 2018;18(1):142. doi:10.1186/s12905-018-0627-8 [PubMed: 30134884]

20. Matteson KA, Scott DM, Raker CA, Clark MA. The menstrual bleeding questionnaire: development and validation of a comprehensive patient-reported outcome instrument for heavy menstrual bleeding. *BJOG Apr 2015;122(5):681–9*. doi:10.1111/1471-0528.13273 [PubMed: 25615842]
21. Facchin F, Barbara G, Saita E, et al. Impact of endometriosis on quality of life and mental health: pelvic pain makes the difference. *J Psychosom Obstet Gynaecol 2015;36(4):135–41*. doi:10.3109/0167482X.2015.1074173 [PubMed: 26328618]
22. Vilos GA, Allaire C, Laberge PY, Leyland N, CONTRIBUTORS S. The management of uterine leiomyomas. *J Obstet Gynaecol Can Feb 2015;37(2):157–178*. doi:10.1016/S1701-2163(15)30338-8 [PubMed: 25767949]
23. Coyne KS, Margolis MK, Murphy J, Spies J. Validation of the UFS-QOL-hysterectomy questionnaire: modifying an existing measure for comparative effectiveness research. *Value Health 2012 Jul-Aug 2012;15(5):674–9*. doi:10.1016/j.jval.2012.03.1387 [PubMed: 22867776]
24. Liddy C, Wiens M, Hogg W. Methods to achieve high interrater reliability in data collection from primary care medical records. *Ann Fam Med 2011 Jan-Feb 2011;9(1):57–62*. doi:10.1370/afm.1195 [PubMed: 21242562]

CONDENSATION

- We present severity indices - incorporating diagnoses, lab and imaging data, and patient-reported symptoms - for the most common indications for hysterectomy.

AJOG at a Glance:**Why was this study conducted?**

- We created symptom severity indices from structured and free text electronic health record data for the most common indications for benign hysterectomy to deepen understanding of drivers of disparity.
- Traditional comorbidity indices have limited utility in younger populations and when outcomes of interest are in treatment choice and not morbidity, mortality, or readmission.
- Hysterectomy, a surgery with striking variation by race and geography, is driven by gynecologic symptom severity, which is currently unmeasured in most population research.

What are the Key findings

- For the three most common indications for hysterectomy – vaginal bleeding, bulk symptoms (pressure), and pelvic pain – groups can now be compared with a measure that incorporates diagnoses, lab values, imaging data, and patient-reported symptoms.
- These composite symptom severity indices were highly associated with clinically relevant objective measures and we were able to identify severe cases not noted by diagnostic codes alone.

What does this study add to what is already known?

- Comprehensive multi-source symptom severity indices can now be used as control variables or proxies for quality of life in health system-derived data in gynecology research.
- Accounting for the degree of symptom severity is particularly important for investigating unexplained inequities in care.

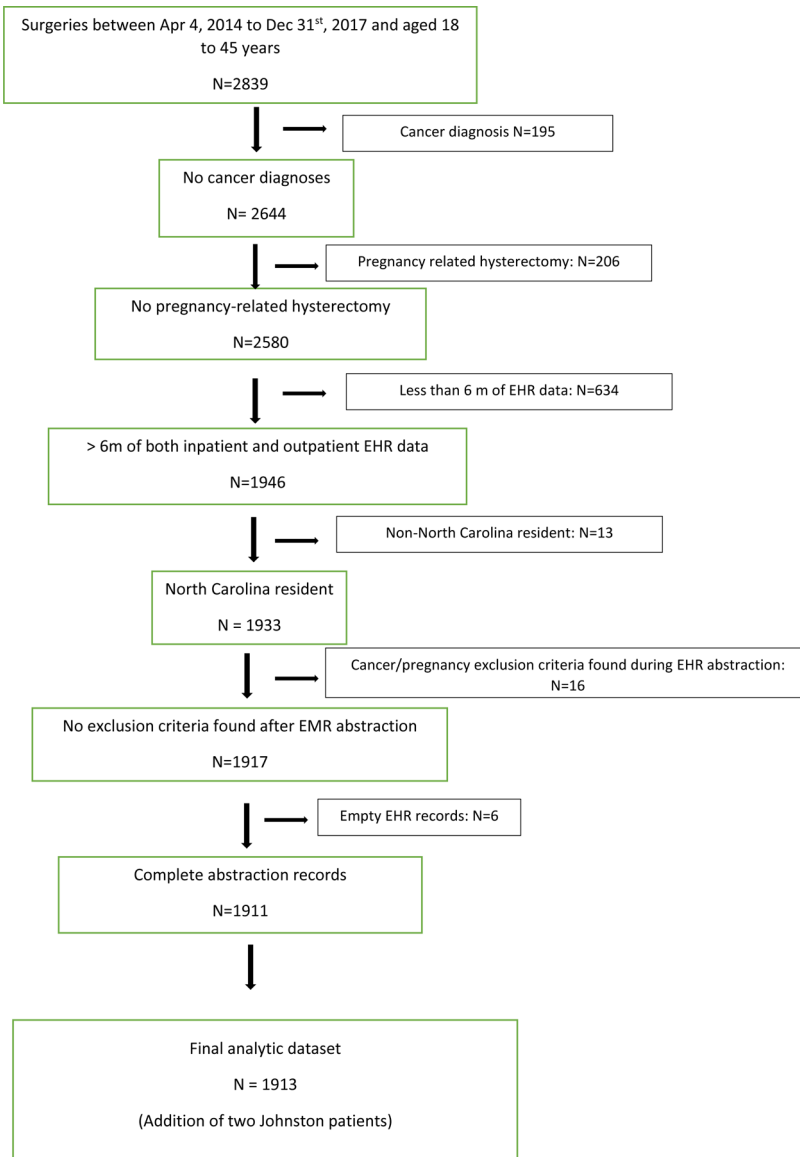


Figure 1: Cohort identification of North Carolina residents treated with hysterectomy for benign, non-emergent indications using electronic health record (EHR) data from a large health care system in the U.S. South, 2014–2017.

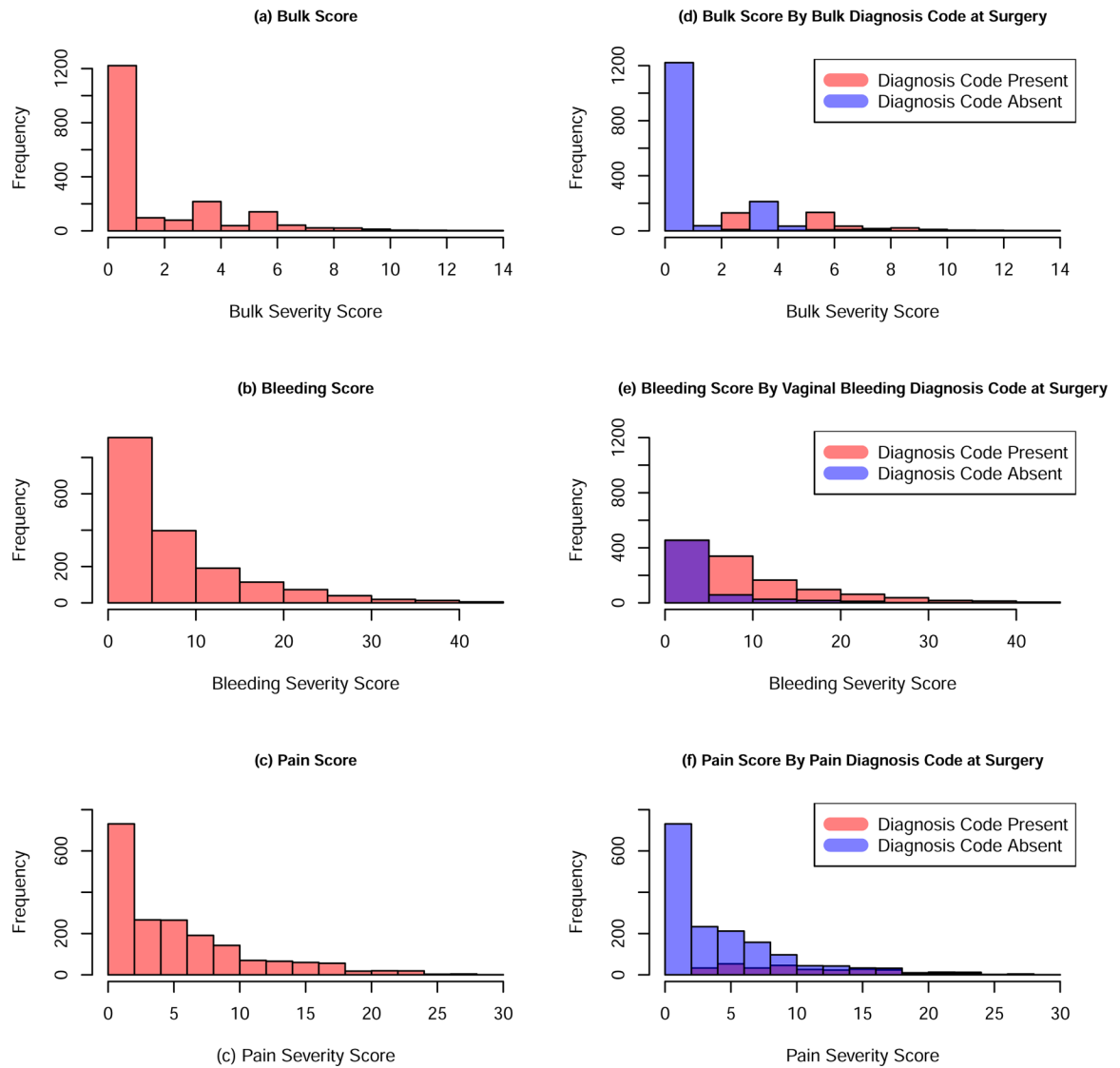


Figure 2: Distribution of Final Composite Symptom Severity Index Scores for Bulk (a), Bleeding (b) and Pain (c) and Final Composite Symptom Severity Index Scores for Bulk (d), Bleeding (e) and Pain (f) Severity Scores stratified by Diagnosis Codes Present (red) or Absent at Surgery (blue). Please note the purple color results from overlapping data.

Table 1:

Descriptive characteristics of individuals between ages 18 and 44 years old treated with hysterectomy in a large not-for-profit health system in the U.S. South, 2014–2017

Variable		
Race/ethnicity – N (%)	Non-Hispanic White	1063 (56%)
	Non-Hispanic African American/Black	580 (30%)
	Non-Hispanic Asian	23 (1%)
	Non-Hispanic American Indian/Alaska Native	17 (1%)
	Hispanic	162 (8%)
	Other	31 (2%)
	Unknown/Refused	37 (2%)
Insurance status – N (%)	Tricare	56 (3%)
	Self-Pay	149 (8%)
	Private Insurance	1,375 (72%)
	Medicare	70 (4%)
	Medicaid	233 (12%)
	Agency	30 (2%)
Hospital type – N (%)	Community	1,047 (55%)
	Rural	5 (0%)
	Teaching	861 (45%)
Year (of surgery) - N (%)	2014	81 (4%)
	2015	479 (25%)
	2016	528 (28%)
	2017	825 (43%)
Age at Hysterectomy – Mean (Range)		39 (19,45)
Uterine size - Mean (Range)		281 (20, 7,031)
Bulk Diagnosis Code [‡] at Surgery (DX) [‡]		366 (19%)
Vaginal Bleeding Diagnosis Code at Surgery (DX)		1,288 (67%)
Pain Diagnosis Code at Surgery (DX)		307 (16%)

[‡]DX: Administrative billing code (ICD9, ICD10, CPT)

[‡]Complete list of all codes can be found in Supplemental Digital Content 1.

Table 2:

Presence of candidate markers of symptom severity in Multi-source data by presence or absence of symptom-specific Diagnostic Code at the time of hysterectomy

Candidate Markers of Symptom Severity in Multi-Source Data	Diagnostic Code Present at Time of Surgery	Diagnostic Code Absent at Time of Surgery
<i>Bulk Symptom Markers (Data Source)</i>	<i>Bulk Diagnostic Codes Present At Surgery (N=366)</i>	<i>Bulk Diagnostic Codes Absent At Surgery (N=1,547)</i>
Bloating (PT) [‡]	15%	7%
Pelvic Pressure (PT)	12%	6%
Uterine Size 50 th -75 th percentile	21%	26%
Non-specified bulk symptoms (PT)	5%	1%
Bulk as Indication for Surgery (MD) [‡]	5%	1%
Uterine size 75 th percentile	60%	17%
<i>Vaginal Bleeding Symptom Markers (Data Source)</i>	<i>Vaginal Bleeding Diagnostic Codes Present At Surgery (N=1,288)</i>	<i>Vaginal Bleeding Diagnostic Codes Absent At Surgery (N=625)</i>
Heavy Bleeding (PT)	69%	31%
Irregular Bleeding (PT)	50%	17%
Heavy Bleeding as Indication for Surgery (MD)	49%	3%
Irregular Bleeding as Indication for Surgery (MD)	35%	5%
Period Lasts Longer than 7 days (PT)	23%	10%
Lethargia or Dizziness (PT)	20%	13%
Iron Use (MD)	33%	16%
1 ER Visit Related to Menorrhagia (DX) [§]	5%	3%
Anemia Diagnostic Code At Surgery (DX)	24%	11%
Anemia as Indication for Surgery (MD)	9%	2%
More than 1 ER Visit Related to Bleeding (DX)	2%	0%
1 ER Visit Related to Anemia (DX)	4%	3%
Anemia Diagnostic Code in Year Prior to Surgery (DX)	19%	9%
History of Blood Transfusion (MD)	7%	3%
More than 1 ER Visit Related to Anemia (DX)	1%	0%
Anemia - HGB < 10 (LAB) [¶]	20%	11%
<i>Pain Symptom Markers (Data Source)</i>	<i>Pain Diagnostic Codes Present At Surgery (N=307)</i>	<i>Pain Diagnostic Codes Absent At Surgery (N=1,606)</i>
Pelvic Pain (PT)	72%	42%
Painful Periods (PT)	51%	30%
Painful Intercourse (PT)	21%	9%
Tylenol (PHARM)	17%	11%
NSAID (PHARM)	35%	30%
Pain as Indication for Surgery (MD)	44%	18%
Painful Periods as Indication for Surgery (MD)	30%	13%

Candidate Markers of Symptom Severity in Multi-Source Data	Diagnostic Code Present at Time of Surgery	Diagnostic Code Absent at Time of Surgery
Other Pain Medication (PHARM) ^{††}	10%	3%
Opioid (PHARM)	41%	31%
At Least 1 Pain Related ER Visit (DX)	17%	8%
Muscle Relaxant (PHARM)	13%	5%

[†]PT: Symptom reported by patient as recorded in unstructured physician notes

[‡]MD: Physician indicated reason for surgery in preoperative or operative notes

[§]DX: Administrative billing code (ICD9, ICD10, CPT)

[¶]LAB: Results from laboratory tests

^{††}PHARM: Prescription information from pharmacy billing data

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

Composite Symptom Severity Index Scoring Method For Bulk, Vaginal Bleeding, and Pelvic Pain

Bulk Severity Index	
Points	Symptoms
1	Bloating (PT) [†] Pelvic Pressure (PT) Uterine Size 50 th -75 th Percentile (LAB)
2	Bulk Diagnosis Code at Surgery (DX) [‡]
3	Non-specified Bulk Symptoms (PT) Bulk as Indication for Surgery (MD) [§] Bulk Diagnosis Code in Year Prior to Surgery (DX)
4	Uterine Size 75 th Percentile (LAB)
Vaginal Bleeding Severity Index	
Points	Symptoms
1	Vaginal Bleeding Diagnosis Code at Surgery (DX) Heavy Bleeding (PT) Irregular Bleeding (PT) Heavy Bleeding as Indication for Surgery (MD)
2	Irregular Bleeding as Indication for Surgery (MD) Vaginal Bleeding Diagnosis Code in Year Prior to Surgery (DX) Period Last Longer than 7 Days (PT) Lethargia or Dizziness (PT)
3	Iron Use (MD) 1 ER Visit Related to Bleeding (DX) Anemia Diagnosis Code at Surgery (DX)
4	Anemia - HGB < 10 (LAB) [¶] Anemia as Indication for Surgery (MD) More than 1 ER Visit Related to Bleeding (DX) 1 ER Visit Related to Anemia (DX) Anemia Diagnosis Code in Year Prior to Surgery (DX)
5	History of Blood Transfusion (MD) More than 1 ER Visit Related to Anemia (DX)
Pain Severity Index	
Points	Symptoms
1	Pelvic Pain (PT) Painful Periods (PT) Painful Intercourse (PT) Tylenol (PHARM) ^{††}
2	NSAID (PHARM) Pain as Indication for Surgery (MD) Painful Periods as Indication for Surgery (MD)
3	Pain Diagnosis Code In Year Prior to Surgery (DX) Pain Diagnosis Code at Surgery (DX) Other Pain Medication (PHARM)
Points	Symptoms
4	Opioid (PHARM) At Least 1 Pain Related ER Visit (DX) Muscle Relaxant (PHARM)

[†]PT: Patient reported symptom as recorded in unstructured physician notes[‡]DX: Administrative billing code (ICD9, ICD10, CPT)

[§]MD: Physician indicated reason for surgery in preoperative or operative notes

[¶]LAB: Results from laboratory tests or pathology report

^{††}PHARM: Prescription information from pharmacy billing data

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript