Development and Validation of Machine Learning Algorithms to Classify Lower Urinary Tract Symptoms

Kai B. Dallas MD[1], Jeffrey N. Chiang PhD[2], Ashley T. Caron BS[3], Jennifer T. Anger MD, MPH[4], Melissa R. Kaufman MD PhD[5], A. Lenore Ackerman MD, PhD[6]

1 Department of Urology, City of Hope National Medical Center, Duarte, CA
2 Department of Computational Medicine, David Geffen School of Medicine at UCLA, Los Angeles, California
3 Michigan State University College of Human Medicine, Grand Rapids, MI
4 Department of Urology, University of California, San Diego, San Diego, California
5 Department of Urology, Vanderbilt University, Nashville, TN
6 Department of Urology, Division of Pelvic Medicine and Reconstructive Surgery, David Geffen School of Medicine at UCLA, Los Angeles, California

Corresponding Author:
A. Lenore Ackerman, MD, PhD
Director of Research, Division of Pelvic Medicine and Reconstructive Surgery
Assistant Professor of Urology and Obstetrics and Gynecology
David Geffen School of Medicine at UCLA
Box 951738 ● Los Angeles, California ● 90095-1738
Phone: (310)794-0206
Fax: (310) 794-0211
email: AAckerman@mednet.ucla.edu

Word Count: 5787
Abstract word count: 419

Conflict of Interest: A.L. Ackerman receives grant funding from Medtronic, Inc. and MicrogenDx, is an advisor for Abbvie and Watershed Medical.

Funding: ALA was supported by NIH K08DK118176.

Abstract

Background: Lower urinary tract urinary symptoms (LUTS), such as urinary urgency, frequency, and incontinence, affect the majority of the population at some point in the lifepspan, causing substantial morbidity, yet few receive effective care. Acccurate diagnosis and treatment is usually dictated by the dominant symptom, however, the sizeable symptomatic overlap between disease categories and subjectivity of language used to describe symptoms leads to high rates of misdiagnosis. We hypothesized that more specific and homogeneous LUTS diagnoses are characterized not by specific pathognomic features, but by patterns of existing symptoms indicative of unique causes of convergent symptomatologies. To improve care and diagnostic acuracy, we sought to employ a data-driven approach to LUTS categorization using machine learning to generate diagnostic groupings based on patient-reported clinical data, creating a novel tool for diagnosis for patients with voiding complaints.

Methods and Findings: Questionnaire responses in a development cohort of 514 female subjects was used for model development, identifying 4 major categories and 9 specific phenotypes of LUTS using agglomerative hierarchical clustering. The dominant features for each cluster and phenotype were examined independently by two urologic specialists and assigned a clinical identity consistent with recognized causes of voiding dysfunction. Then, a supervised machine learning model was trained to assign unseen patients into these phenotypes. The model was then applied to a validation population of 571 additional subjects to validate the diagnostic machine learning algorithm, demonstrating good reproducibillty of the phenotypes and their symptomatic patterns in an independent cohort. This data-driven, hierarchical clustering approach captured overlapping symptoms inherent in typical patients, recognizing common uncomplicated diagnoses (e.g. overactive bladder) and several novel diagnostic categories (e.g. myofascial pelvic pain). The assigned clusters and phenotypes were consistent with coded primary diagnoses with 70% accuracy; however, physician primary diagnosis may not reflect the complete range of patient symptoms as patients often present with multiple urinary symptoms that do not fit pre-established diagnoses. Although this analysis reflects care-seeking populations from a single center, further studies will need to determine if the model is scalable to the population at large.

Conclusions: We describe the generation of a machine learning algorithm relying only on validated patient-reported symptoms for accurate diagnostic classification. Given a growing physician shortage and increasing challenges for patients accessing specialist care, this type of digital technology holds great potential to improve the recognition, diagnosis, and treatment of functional urologic conditions. While future prospective work with larger, multi-institutional cohorts is needed, with refinement, this approach is capable of increasing both the equity and rapidity of access to effective urologic care.

Introduction:

Half of adult women are incontinent; more than 75% of women at some point in the lifespan report at least one lower urinary tract symptom (LUTS), including *urinary urgency, frequency, nocturia, painful urination, bladder pain or discomfort, or incontinence,* that disrupts daily activities.(1, 2) When severe, these urinary symptoms degrade health-related quality of life to levels worse than chronic dialysis.(3) Incontinence and urinary complaints are one of the most common causes for the long-term institutionalization of older adults.(4) Further, these disorders represent substantial economic burden, with annual costs estimated to approach 100 billion dollars.(5) Yet despite this large burden of illness, most women do not receive care for these symptoms. For those who do seek care, many will not receive an accurate diagnosis.(6) And for the few who do obtain care in specialty urology clinics, more than 90% will abandon medical therapies within a year.(7)

Numerous challenges to effective treatment face patients seeking care for LUTS. While patients presenting with urinary symptoms are categorized into several unique conditions, appropriate diagnosis and treatment is complicated by subjectivity and sizeable symptomatic overlap.(8) Conditions such as overactive bladder (OAB) and interstitial cystitis/painful bladder syndrome (IC/BPS) are symptom complexes, not diseases, that likely represent multiple pathophysiologies requiring different treatments.(13, 15, 16) Few therapies for these conditions are widely effective; high non-responder rates are likely indicative of unrecognized, physiologically distinct phenotypes. Despite being considered distinct entities, the presentations of OAB and IC/BPS can be similar.(9, 10) Similarly, the causes of symptoms can be unclear. It can be challenging to distinguish when symptoms of urinary urgency result from OAB or a cystocele, particularly when the patient has no other symptoms attributable to pelvic organ prolapse (POP).(11, 12) As there are currently no objective tests or biomarkers available to make such distinctions, diagnosis and treatment are based entirely on clinician judgement.(13, 14) The first step in improving the care delivered to these patients is refining the diagnostic classification of patients into more specific phenotypes.

Two decades ago, a leader in urology charged the field to achieve this goal by developing algorithms to overcome barriers to care delivery and promote novel, cost-effective approaches to personalize therapy.(17) Although the last decades have not seen progress toward such implementation, current digital technology may have put this goal in reach. Machine learning (ML) represents a body of analytical methods with promise in assisting clinicians in disease diagnosis across fields.(18-22) ML has been applied to urologic diseases, such as urolithiasis, bladder cancer, urinary tract infection, and kidney cancer.(23-28) In voiding dysfunction, a study using a random forest-based ML algorithm successfully predicted the success of anticholinergics for OAB with 80% accuracy.(29) ML has also been applied to the study of voiding pathways in patients with multiple sclerosis.(30)

A data-driven approach to identifying more specific phenotypes may recognize subtle distinctions indicative of unique causes of convergent symptomatologies. While symptoms of LUTS diagnoses may be similar, the underlying etiologies of these symptoms may be quite different and require different approaches to treatment. To this end, multiple studies have attempted to utilize a range of unsupervised ML methods for better phenotypic classification of lower urinary tract disorders. However, most have examined only a small subset of symptoms, typically excluding patients with overlapping symptoms or confounding factors, such as POP. For most ML methods, data objects are divided into non-overlapping, mutually-exclusive clusters of a fixed number. In contrast, hierarchical clustering generates a set of nested clusters

that are arranged as a tree/dendrogram. Hierarchical clustering is best used when the number of classes is not known; once the structure of the data is defined, one can stop at any number of clusters. A strength of hierarchical clustering is the ability for the resulting dendrogram to provide insights as to the structure of the data, which in turn informs about the number of logical clusters. Given the nature of functional urologic symptoms, we believe that hierarchical clustering presents a more complete picture of the underlying patterns of lower urinary tract symptoms than other methods of machine learning.

Given the known clinical dilemmas in diagnosing LUTS, we aimed to create an unsupervised ML algorithm encompassing all patients presenting with LUTS. This novel tool for the diagnosis of patients with voiding complaints utilized agglomerative hierarchical clustering to generate diagnostic groupings based on patient-reported clinical data.

Methods:

Study Cohorts

After Institutional Review Board (IRB#00040261) approval, the development dataset of 514 female subjects with a wide range of symptoms, seen consecutively for evaluation in a tertiary urogynecology practice between January and December 2018, was used for model development. A second dataset of 571 subjects recruited from the same practice consequtively between January 2019 and March 2020 was used to validate the diagnostic ML algorithm. For each cohort, subjects were administered four validated questionnaires at initial evaluation: the female Genitourinary Pain Index (fGUPI)(31), Overactive Bladder Questionnaire (OABq)(32), Pelvic Floor Distress Inventory (PFDI-20)(33) and O'Leary−Sant Indices, including the Interstitial Cystitis Symptom and Problem Indices (ICSI/ICPI).(34) The fGUPI measures the nature and severity of genitourinary pain, and contains subscales assessing pain, urinary symptoms, and quality of life.(31) The OABq measures continent and incontinent OAB symptoms, containing subscales for symptom bother, coping behaviors, concern/worry, social interaction, sleep, and health-related quality of life.(32) Only the symptom-based questions (1-8) were utilized in this analysis. The PFDI-20 measures pelvic floor symptoms, including domains that assess urinary, defecatory, and prolapse symptoms.(33) The ICSI/ICPI are used together to measure the severity and bother of urinary frequency, urgency, nocturia, and bladder pain.(34) Patients with active urinary tract infection (UTI), prior pelvic reconstructive surgeries (such as prolapse repairs, bladder augmentation, incontinence procedures), current pregnancy, or neurogenic bladder were excluded from the study. Patients with cyclic pain at menses were excluded, however, patients with comorbid functional pain syndromes, such as irritable bowel syndrome or fibromyalgia, were allowed to participate. Patients with asymptomatic complaints (e.g., microscopic hematuria) were included in both datasets.

Hierarchical Clustering

Agglomerative hierarchical clustering was applied to the development dataset to identify patient groups on the basis of patient age and questionnaire responses. This unsupervised approach recursively builds a hierarchical representation from the "bottom up" by grouping pairs of samples according to a distance metric and linkage criterion. This algorithm was applied using the Ward linkage criterion, which aims to create groupings which have minimum intra-group variance.(35) Similarity between samples was calculated using the Euclidean (i.e., squared) distance between symptoms. After the hierarchical representation was constructed using the agglomerative algorithm, clusters were created by truncating the depth of the hierarchical representation, thereby grouping patients into a fixed number of clusters (e.g., **Fig. 1**). Mean silhouette scores(36) were computed to measure cluster assignment performance to select the optimal cluster assignment between 2 and 15 clusters. Using this criterion, four and nine cluster solutions were selected (see Results).

Phenotypic group descriptive assignments

Age and survey response scores were scaled to a set range before plotting as a heat map. ANOVA assessed significant differences between cluster groups in intragroup means for each variable. The dominant features for each cluster and phenotype were examined independently by two urogynecologists, who each assigned a clinical identity based on their expert opinion. For any disagreement, discussion between the two evaluators and a third moderating urogynecologist allowed for refinement of the group clinical identity until consensus was reached.

Generation of the diagnostic algorithm

Unsupervised algorithms, such as agglomerative clustering, are not reliably applicable to new data as they require the entire hierarchical representation to be re-generated each time the algorithm is run. After manually validating cluster assignments and assigning phenotypic group descriptions, a supervised machine learning model was trained to assign unseen patients into existing clusters. Using the development dataset, random forest models were trained to predict the machine-generated cluster assignment on the same symptoms used for clustering. Random forests are generally robust to overfitting and outliers, and outperform other methods when representing tabular data.{Grinsztajn, 2022 #1578} Two models were generated which assigned patients into the four and nine cluster labels, respectively. These models were assessed using the validation dataset with balanced accuracy(38) (defined as the average sensitivity score for each class) – which has been used to deal with imbalanced data.

Uniform manifold approximation and projection (UMAP) visualization
UMAP is a nonlinear dimension reduction technique used to visualize high-dimensionality data for qualitative inspection.(39) Using the development dataset, the algorithm was applied to represent patients in two data-generated axes. As in the clustering analysis, the projection axes were optimized with respect to the Euclidean distance between patient responses. Then, validation dataset patients were transformed into the same space and all patients visualized together. Each patient was colored according to the phenotype assigned by the nine-cluster model.

Accuracy of cluster assignment
We determined how often the phenotypic cluster was consistent with coded diagnosis. Primary and secondary ICD-10 codes affiliated with the consultation at which symptoms were assessed were catalogued. The 10 most common clinical diagnoses were aligned with one or more ML cluster(s) and phenotype(s) (Supplemental Table 2) by two FPMRS specialists blinded to the other physician's assessment and the clustering groups (clinician cluster groups). A third specialist ajudicated any disagreements.

Statistical analysis
Cluster fitting and optimization, machine learning model development and evaluation, and UMAP were performed with Python using the scikit-learn and UMAP-learn packages. All other analyses and visualization were performed in R version 4.2.1. Differences in demographic and clinical characteristics were compared by using Wilcoxon signed rank tests for paired data and the Pearson chi square, Fisher exact, or Mann-Whitney U-tests for independent data as appropriate (2-tailed). Differences in proportions were compared using the two-sample z-test. Results were considered significant at an alpha level <0.05.

Results

Unsupervised clustering of subjects presenting with LUTS.
The development dataset of 514 female patients with a mean age of 58.7 years completed a panel of validated questionnaires between June 2017 and December 2018. Using only age and patient-reported symptoms, subjects were classified into symptomatic clusters according to an agglomerative hierarchical clustering algorithm (**Fig. 1**). Examination of the resulting dendrogram (**Fig. 1**, blue line) revealed four clear branches, symbolic of symptomatic patterns for patients seeking urogynecologic care, confirmed as the optimal number of clusters using the silhouette method (global maximum, **Fig. 2**).

Data-driven hierarchical clustering reveals the global patterns in phenotypes.
The symptom patterns associated with each cluster (**Table 1**) reflected the overall categories of subjects seeking urogynecologic care: 1) subjects with mild symptoms who are only moderately bothered (*Mild symptoms/controls*), 2) subjects with predominantly lower urinary tract symptoms, such as urinary frequency, urgency, and incontinence (*LUTS*), 3) subjects with pain as their predominant complaint (*Pain*), and 4) subjects with pelvic floor complaints (pelvic pressure, vaginal bulge) who also exhibit urinary symptoms and discomfort (pelvic floor disorders [*PFD*]). These patterns are more clearly visualized in a heat map (**Fig. 3**) demonstrating which features are prominent in each group. The *PFD* group had globally elevated symptom scores across all symptomatic domains, with unique complaints attributable to the pelvic floor, such as defecatory dysfunction (PFDI20 q7-8), vaginal bulge (PFDI20 q3), pelvic pressure (PFDI20 q2), and a need to splint to defecate or void (PFDI20 q4, PFDI20 q6). The *Pain* group was the youngest group overall and displayed similarly high scores to the *PFD* group only on questions assessing pelvic and genitourinary pain (fGUPI q1-4, PFDI20 q20, ICSI q4, ICPI q4). The *LUTS* groups was generally older than the other groups and exhibited elevated urinary incontinence, both urge (PFDIU20 q16, OABq q8) and stress (OABq q4, PFDI20 q17), occurring in concert with urinary urgency (OABq q3, ICPI q3, ICSI q1), frequency (ICSI q2, ICPI q1, GUPI q6), and nocturia (ICPI q2, OABq q5). The symptom severities of the *LUTS* group surpassed the *PFD* group only in the domain of fecal incontinence (PFDI20 q9-11).

Discrimination of symptomatic phenotypes.
While the four-cluster classification provides a general categorization of women presenting for urogynecologic care, the resulting groups encapsulate only the most basic symptomatic categorization, failing to distinguish patients for whom different treatments are selected, such as stress and urgency incontinence, vestibulodynia and interstitial cystitis/bladder pain syndrome, or pelvic organ prolapse and myofascial pelvic pain. We therefore examined the next most optimal range of clusters. The silhouette score reached a local maximum in the range of 9-11 clusters (**Fig. 2**, local maximum); in the dendrogram, the nine-cluster division best captured the major subdivisions in the data (**Fig. 1**, gold line). Again visualized as a heat map (**Fig. 4**), patterns of dominant symptoms provide characterization of each cluster phenotype (**Table 2**).

Mild symptoms: Subjects in the *Mild symptoms/controls* cluster were notable primarily for low bother scores (fGUPI q7-9), but further subdivision of this group isolated two additional phenotypes. The first was comprised of subjects with mild symptoms and who endorsed very little bother. While this subgroup exhibited a range of mild urinary symptoms, SUI was the most common cause for seeking care, so we dubbed this group the *Control/SUI* phenotype. A second phenotype, denoted Non-urologic Pelvic Pain (*NUPP*), complained of painful symptoms localized to the vagina and urethra (fGUPI q1a, fGUPI q1c), often manifesting as dyspareunia (fGUPI q2b), but lacking bladder pain or pain with bladder filling (fGUPI q2c-d).

LUTS symptoms: The *LUTS* cluster was further classified into three distinct phenotypes, each of which demonstrated some urinary urgency and urgency urinary incontinence (UUI): 1) subjects with mild-moderate urgency and UUI (OABq q8, PDFI-20 q16) with or without stress incontinence (*UUI*), 2) subjects with more severe urgency, frequency, nocturia, and UUI (ICSI q1, ICPI q3, OABq q8, PDFI-20 q15-16), who also exhibited severe stress urinary incontinence (PFDI-20 q17) [mixed urinary incontinence (*MUI*)], and 3) subjects with MUI and nocturia (ICPI q2, OABq q5) who also endorsed fecal urgency (PFDI-20 q13) and incontinence (FI) (PFDI-20 9-11) (*UUI/FI*).

Pain symptoms: The Pain cluster was divided into two phenotypes, the first of which was unique in the severity of endorsed bladder pain and discomfort (fGUPI q4, ICPI q4, PFDI-20 q20), features classically associated with interstitial cystitis/bladder pain syndrome (*IC/BPS*). The second phenotype also endorsed pain, although milder, but were singular in their endorsement of a sensation of incomplete bladder emptying (fGUPI q5, PFDI-20 q5). As the occurrence of this specific symptom in the absence of prolapse is commonly associated with a high-tone or non-relaxing pelvic floor, we designated this group high-tone pelvic floor dysfunction (*HT PFD*).

PFD symptoms: While both phenotypes derived from the PFD cluster exhibited a wide range of urinary, painful, and pelvic floor complaints, the type and severity differed between the two phenotypes. One phenotype endorsed vaginal bulge (PFDI-20 q3) or the need to splint while voiding or defecating (PFDI-20 q4, PFDI-20 q6), symptoms highly specific for pelvic organ prolapse (*POP*). In contrast, the second group had more severe, uncomfortable urinary symptoms (e.g., OABq q2) and pelvic pain (fGUPI q4) occurring in a similar pattern to that seen for the *HT PFD* phenotype, which we denoted as myofascial pelvic pain (*MPP*).

Development of a classification algorithm

We next sought to create a diagnostic tool for the phenotypic classification of new patients. We trained a random forest classifier using the phenotypic assignments of the development dataset to create a classification algorithm, which we then applied to the independent validation dataset of 571 individuals with urogynecologic complaints. Phenotypic assignment using the random forest classifier resulted in groups that were highly similar to the development dataset phenotypes, exhibiting the same patterns of distinctive symptoms (**Supplemental Table 1**). We plotted the similarity of the development and validation phenotypes graphically using a uniform manifold approximation and projection (UMAP) method (**Fig. 5**). Proximity in UMAP space signifies similarity between groups, in this case capturing the overlap in symptom profiles between phenotypes from the validation and development cohorts as well as the relationships between groups. For example, the MUI group exists at the interface between the UUI and SUI groups, sharing features of both while still occupying a unique space. Similarly, the MPP group is situated between the IC/BPS, HT PFD, and POP phenotypes, as it shares a subset of symptoms with each, but also exhibits specific symptoms that make the phenotype distinct.

Accuracy of cluster assignment.

We next examined how well these assigned phenotypes aligned with coded diagnoses for the validation cohort. As many of the phenotypic designations do not have associated diagnostic codes in the ICD10 diagnostic classification system, we first aligned the 10 most common diagnostic codes with the most closely related phenotypic clusters for both the four- and nine-cluster solutions (**Supplemental Table 2**). For example, overactive bladder is a diagnosis characterized predominantly by urgency and urgency incontinence, thus overactive bladder subjects were considered accurately classified if assigned to the LUTS cluster or the UUI, MUI, or FI/UUI phenotypes. Using this approach, the assigned clusters and phenotypes were

consistent with diagnoses with 71% and 70% accuracy for the four- and nine-cluster models, respectively (**Fig. 6**).

Further subdivision of the clusters may be meaningful.
While subdivision of the population into nine symptomatic phenotypes created a rational classification system consistent with observed diagnoses, further subdivision may still prove meaningful. Following the branch of the dendrogram encompassing the NUPP group, there is a clear downstream split of this patient cluster into two additional groups (**Supplemental Fig. 1**). The primary difference between these subclusters was age (70±7 vs. 35±10 years of age, $p < 0.001$). While both groups exhibited vaginal, introital, and urethral pain, the proportions of older patients with each painful complaint was higher, but the severity of pain was higher in younger patients. In addition, the older patients had more nocturia (ICPI q2, OABq q5) and urinary urgency (OABq q3), while the younger patients exhibited more urinary frequency (ICSI q2, ICPI q1).

Similarly, within the cluster of patients with mixed incontinence (*MUI*) was a subject subgroup endorsing complaints consistent with concomitant anterior vaginal wall prolapse. While this subgroup (*Cystocele/MUI*) specifically endorsed significant bulge (PFDI20 q3), pelvic pressure (PFDI20 q1) and heaviness (PFDI20 q2), these subjects otherwise had similar levels of SUI, urinary frequency, and nocturia to the *MUI* phenotype (**Supplemental Table 3**). Despite these symptomatic hallmarks of prolapse, the *Cystocele/MUI* subjects were significantly different from the *POP* group, lacking the defecatory complaints (PFDI20 q7-14) and splinting (PFDI20 q4, PFDI20 q6) seen for the *POP* phenotype.

Discussion

We successfully applied machine learning algorithms to the diagnostic classification of women with a wide range of symptoms presenting for urologic care. This classification generated logical, phenotypic groups based on validated patient-reported symptoms alone. Symptomatic patterns could be grouped into four general clusters: minimal/mild symptoms, urogenital pain, urinary complaints, and pelvic floor disorders. Validation of these clusters revealed high reproducibility in an independent cohort with an accuracy of 71%. These groups are analogous to the general clinical categories currently used: most patients presenting to urogynecology clinics will be diagnosed with either incontinence, genitourinary pain, or pelvic organ prolapse. As unsupervised machine learning brings no assumptions to cluster derivation, agreement of the overall diagnostic categories with well-accepted clinical categorization validates the ability of data-driven methods to derive clinically meaningful diagnostic categories.

There are, however, limitations to this simplistic categorization. Conditions such as OAB and IC/BPS are symptom complexes, not diseases, that likely represent multiple pathophysiologies requiring different treatments.[13,37,38] Few therapies for these conditions are widely effective; high non-responder rates are likely indicative of unrecognized, physiologically distinct phenotypes. The first step in improving the care delivered to these patients is refining the diagnostic classification of patients into more specific phenotypes.

Several groups have tried to subclassify urinary symptoms or genitourinary pain, but have typically examined only one symptom cluster in isolation (e.g., overactive bladder(40-43) or genitourinary pain(44-47)). However, “as often patients present with multiple urinary symptoms that do not perfectly fit the pre-established diagnoses,”(42) the intentional exclusion in these studies of patients with overlapping symptoms or co-existing pelvic organ prolapse may underlie the lack of practical progress in improving current diagnostic schema. Prior work examining the clinical overlap between OAB and IC/BPS noted many patients carrying an OAB diagnosis suffer from significant bladder pain (35%) while 35% of IC/BPS patients have urge incontinence.(9, 10) In addition, OAB symptoms are common in patients desiring surgical repairs for POP,(12) primarily apical and anterior prolapse. Given the high prevalence of prolapse in the population (24-49%),(48-50) it is unclear how co-existing POP should influence our clinical management of patients presenting primarily with urinary complaints.

To overcome this obstacle, broad inclusion of all patients consecutively presenting for urogynecologic care combined with unsupervised clustering using patient complaints alone allowed us to derive nine unique phenotypes encompassing the range of overlapping symptoms without bias. Distinction between groups was based on unique combinations of symptoms rather than individual, pathognomonic features.(41) The resulting groups captured the ranges of coexisting symptoms while still accounting for the complicated symptomatic overlap of real-world patients.

These diagnostic groups were concordant with the coded diagnoses for patients 70% of the time. Some of the inaccuracies may represent suboptimal performance of the algorithm; further cluster subdivision may be needed to distinguish functionally different groups. As discussed above, the *Cystocele/MUI* phenotype, patients presenting primarily with urinary complaints who also had symptoms of anterior prolapse without obstructed voiding (e.g., splinting, difficulty emptying) were distinguishable from other patients with UUI using hierarchical clustering. But these subjects were much more similar to subjects with UUI than subjects with more severe POP symptoms (splinting to defecate or urinate). The management of these *Cystocele/MUI* subjects is a subject of extensive debate. While patients with high-grade prolapse who also

complain of UUI may be cured by surgical prolapse repair,(12) in patients with less severe prolapse it is unclear whether UUI symptoms should be independently managed first or correction of the prolapse attempted in the hopes of managing the urinary symptoms even though prolapse is not the patient's primary complaint. The ability of the algorithm to distinguish these less severe prolapse patients from those with symptoms more clearly attributable to prolapse (*POP* phenotype) may help answer this question. The flexibility of hierarchical clustering allows additional insights and refinement of classification with additional algorithmic training, particularly after validating the results with meaningful outcomes, such as treatment responses.

Subjects with OAB, by far the largest diagnostic group, were subclassified in our analysis into three phenotypes that are well accepted causes for care-seeking: 1) urgency-predominant symptoms and incontinence (*UUI*), 2) mixed incontinence (*MUI*) and 3) mixed incontinence with FI (*FI/UUI*). This third group in particular highlights the utility of ML and digital technology. While FI is estimated to affect 7-15% of community dwelling adults, most patients do not volunteer the symptoms and many providers do not screen for it.(51, 52) This algorithm identified severe FI in a small subset of patients, not one of whom carried FI as a coded diagnosis. In this manner, algorithmic classification may drastically improve the recognition of stigmatizing symptoms that patients may be reticent to divulge in person. Thus, while the technical accuracy of diagnosis for this subgroup was poor, we suspect this reflects a failure of the physician to recognize and affiliate the code for FI as a diagnosis, not a true failure of the algorithm.

Outside of FI, the worst discordance between coded diagnosis and phenotypic classification was for patients diagnosed with POP, but this discrepancy may provide insight into the overlap of pelvic floor symptoms with LUTS. While most diagnostic schema do not address pelvic floor myofascial dysfunction as an independent contributor to urinary symptomatology, the pelvic floor plays a central role in pelvic visceral function and, when dysfunctional, can cause pelvic pain syndromes and urinary complaints.(53) While muscle laxity is acknowledged as a contributor to POP or SUI, muscle hypertonicity or discoordination is rarely recognized. While recognized as a contributor to dysfunctional voiding, dyssynergic engagement of the pelvic floor exhibits a vast spectrum of clinical presentations and widely fluctuating estimates of its prevalence which are indicative of a high rate of misdiagnosis and poor understanding of the contribution of the pelvic floor to urogynecologic complaints.(54) A growing body of evidence, however, suggests that increased pelvic floor tone underlies many urinary, gastrointestinal, and sexual complaints, even in the absence of pain,{Voorham-van der Zalm, 2008 #1520;Voorham-van der Zalm, 2008 #1520} although this supposition is not universally accepted. Our phenotypes identify a subset of subjects with a range of urinary and painful complaints attributable to myofascial dysfunction, embodied in the myofascial pelvic pain (*MPP*) and high-tone pelvic floor dysfunction (*HT PFD*) phenotypes.

In a large study of over 800 patients with pelvic floor myofascial pelvic pain (MPP), MPP subjects endorsed subjective symptoms of prolapse (pelvic pressure and heavinesss) without the objective symptoms of prolapse (bulge, splinting) as well as urinary symptoms (sensation of incomplete emptying) and defecatory dysfunction.(56) In our study, the algorithm divided patients with pelvic floor complaints into a group with clear prolapse symptoms (bulge, splinting) and subjects correlating with the described MPP phenotype. And while this group has elevated pain in comparison to subjects with true prolapse (fGUPI q4, **Fig. 4**), the pattern of symptoms is highly similar between POP and MPP phenotypes. Thus, the MPP group presents like "prolapse without prolapse".(56) Despite a similar clinical presentation to POP, however, MPP can be distinguished by the algorithm from true POP with good reliability, even without a physical exam. These MPP subjects had also been distinguished from POP subjects by urogynecologic

specialists, who most commonly conferred diagnoses of either pelvic pain or pelvic floor dysfunction, reinforcing both the nature of this condition and the ability of the algorithm to distinguish these patients from symptoms alone. Unfortunately, this common clinical presentation of MPP is sparsely documented in the literature, leading to poor awareness of this condition. MPP is the predominant underlying cause in 60-90% of women with chronic pelvic pain,(57-59) which in turn affects an estimated 26% of the world's female population.(60, 61) Taken together, these estimates suggest about 15-20% of adult U.S. women suffer from MPP. It is the primary reason for seeking care in approximately 10% of all gynecologic and urologic visits,(62, 63) which does not account for additional presentations to primary care, gastroenterology, and pain management specialists. Given the scale of this public health issue, the potential utility of the algorithm in identifying *MPP*, particularly without the need for evaluation by a subspecialist skilled in discriminate pelvic exam, could provide better recognition of the condition with earlier assignment to effective care.

The algorithm identified a second myofascial group, *HT PFD*, whose most notable symptoms were urinary in nature: a dominant sensation of incomplete bladder emptying in concert with urinary frequency and urgency prompted more by discomfort than fear of incontinence. While these patients did have some pain, the severity was lower than that seen in other pain conditions (MPP, NUPP, IC/BPS) or POP patients. It is well recognized in the colorectal literature that a high-tone or hypertonic pelvic floor can lead to symptoms of fecal urgency, frequency, and obstructive defecation;(64) normal urinary function similarly requires appropriate relaxation and coordination of the pelvic floor muscles and sphincters. Thus, a growing body of evidence suggests that a high-tone pelvic floor ("non-relaxing" or fixed) can result in analogous symptoms in the urinary tract, even in the absence of pain.(65-69) The algorithm's *HT PFD* phenotype appears to correlate with this condition, exhibiting frequency, uncomfortable urge, and a sensation of incomplete emptying.(70) Despite being common,(68) HT PFD is often mistaken for IC/BPS, OAB, POP, or even UTI, due to the vague mix of uncomfortable bladder sensations and urinary complaints.(68, 69, 71) No specific ICD-10 code specifies this diagnosis, again leading to infrequent recognition and poor treatment outcomes. The majority of HT PFD patients were assigned the M99.05 ICD-10 code generally used to specify pelvic floor dysfunction; this result indicates not only that many urogynecologists are aware of and able to recognize HT PFD but that the algorithm could effectively distinguish this condition from confusable conditions, again without an exam. Given the non-specific constellation of symptoms, it can be challenging for non-specialty providers to recognize HT PFD, which relies on a detailed pelvic exam, good understanding of the global symptom complex, and ruling out confounding conditions.(69) With algorithmic identification of these patients, the correct diagnosis can be made and earlier interventions with the correct treatment (pelvic floor physical therapy) achieved, even when providers are unfamiliar or uncomfortable making the diagnosis.

Almost all of the patients diagnosed with some form of pelvic pain (including IC/BPS, pelvic pain, and urethral pain) were contained within the *Pain* cluster. Outside of IC/BPS, however, the additional pain phenotypes recognized by our algorithm (NUPP, MPP) do not have diagnostic code correlates and can only be recognized symptomatically. However, for those patients who were given more specific symptomatic codes (e.g., vaginal or urethral pain), these symptomatic descriptions aligned well with the phenotypic categories, reinforcing the concept that our current diagnostic approaches lack the necessary distinctions to effectively categorize the spectrum of painful genitourinary complaints observed clinically. Interestingly, the results of this phenotyping were quite similar to a separate analysis in a non-overlapping population of patients with pelvic pain(47) which identified the same three populations of patients with perceived bladder pain (NUPP, MPP, and IC/BPS). While those populations are not well distinguished by clinicians, we

previously demonstrated differences in responses to different treatment approaches, also supporting the idea that these distinct groups are clinically meaningful.

As alluded to above, a challenge in interpreting the accuracy of clustering assignments is that it requires complete and accurate clinician assignment of the patient's dominant complaint as the primary diagnosis. Many patients had secondary diagnoses that covered different diagnostic categories, reinforcing the pervasive diagnostic confusion of patients presenting with urogynecologic complaints as well as likely insufficiency of current coding systems. We are measuring accuracy against a diagnostic coding system that we know lacks the detail needed to describe the range of conditions experienced by our patients. Given these inadequacies, the 70% accuracy of the nine phenotype model likely underestimates practical accuracy in the clinical setting. Although this diagnostic system will need further refinement and validation, this objective, data-driven approach to patient classification may be more meaningful than the current, highly subjective approaches to diagnosis that continue to limit our progress towards improving urogynecologic care.

While we describe subdivision of the cohort into nine phenotypes, we have seen that more subcategorization may be useful. Hierarchical clustering provides a flexible mechanism to adapt with clinical validation; incorporation of additional data, such as treatment outcomes, can continue to train the algorithm to determine if further subdivision is clinically meaningful. As with the NUPP group, further clustering divided patients by age, but it is unclear as yet whether this distinction provides additional clinical guidance. However, despite similar symptoms across the two subgroups, differences in age may belie differences in pathophysiology. The most common cause of vaginal, introital, and urethral pain in older women is vaginal atrophy, a result of hypoestrogenism best treated with local estrogen replacement. As this is unlikely to be the cause of symptoms in a younger population, distinction of younger patients with NUPP from other forms of pelvic pain may allow better exploration of the underlying causes of pain and development of treatments tailored to this specific population.

Our findings here are novel and timely with the potential to improve diagnosis and treatment and overcome the significant burden these conditions place on patients and the healthcare system. While multiple groups have proposed various approaches to subclassifying urinary symptoms and genitourinary pain,(40-46) most require detailed information (patient demographics, physical exam findings, imaging, genetic or biochemical markers, or other diagnostic testing results) unavailable or unfamiliar to most practioners outside of specialized clinical settings. In addition, most of these categorization schema example narrow populations without overlapping symptoms, which has ignored a large subset of real-world patients. Here, we generate a diagnostic algorithm based on our novel ML-based phenotypic classification that can be used for treatment assignment of a broad care-seeking population with a wide range of urinary and pelvic complaints. By relying only on patient-reported symptoms, this algorithm can be applied by any type of practitioner in any care delivery setting, including telemedicine.

Providing adequate care for the enormous population of women affected by urinary complaints faces numerous obstacles, both at the health system and individual levels. First is the lack of recognition of these symptoms. Patients, even when highly bothered, may not seek care for urinary complaints due to shame, embarrassment, or a feeling that symptoms are unavoidable. Doctors may not ask about these issues, and even if they do, may not be able to address them due to time constraints or knowledge gaps. The prevalence of urinary symptoms recognized in the primary care setting increases dramatically when validated questionnaires are routinely administered,(72) stressing the need for better, more pervasive, and perhaps more anonymized screening. As a result of these numerous barriers, however, few women seek or receive care.(73)

The development of a digital resource providing diagnostic assistance has the potential to dramatically improve the quality of life of the population, not only through more accurate diagnostics but by allowing rapid implementation of first-line, low-risk treatments and earlier access to specialty care.

Although our results are promising, our algorithm is based on data from a single center; thus, our findings may not be scalable to the population at large. Further, accurate classification of patients is only of clinical value if the treatment based on these diagnostic groupings correlates with clinical improvement, which we hope to address in future prospective studies. Also, the current modeling requires patients to answer nearly 50 questions, which may not be achievable in most clinical scenarios. Finally, our approach to classifying patients may not be generalizable to populations in different clinical scenarios, such as those not presenting for urogynecologic care.

Despite these limitations, this study examined a large number of patients (over 500 in each cohort, totaling over 1,000 subjects) who were consecutively included regardless of referral diagnosis or presenting complaints. This means that our cohort not only included a control group, but is likely representative of the true patient population referred to urogynecologic practice. Furthermore, this algorithm relies only on validated questionnaire responses. Thus, this novel LUTS classification algorithm can be utilized to assign treatment plans without the need for either sub-specialist evaluation, to which access can be limited, or physical examination, which can be challenging for patients in underserved areas. Given a growing physician shortage and increasing challenges for patients accessing specialist care, digital technology holds great potential to improve the early recognition, diagnosis, and early treatment of functional urologic conditions. Future prospective work with larger, multi-institutional cohorts are needed to improve these algorithms to allow accurate diagnosis and treatment assignment based on the machine learning-suggested phenotypes. In addition, reduction of the number of questions to the minimum subset possible is needed to make this feasible to perform in a clinical setting. With refinement, however, this approach to care delivery is capable of increasing both the equality and rapidity of access to effective urologic care.

Figures

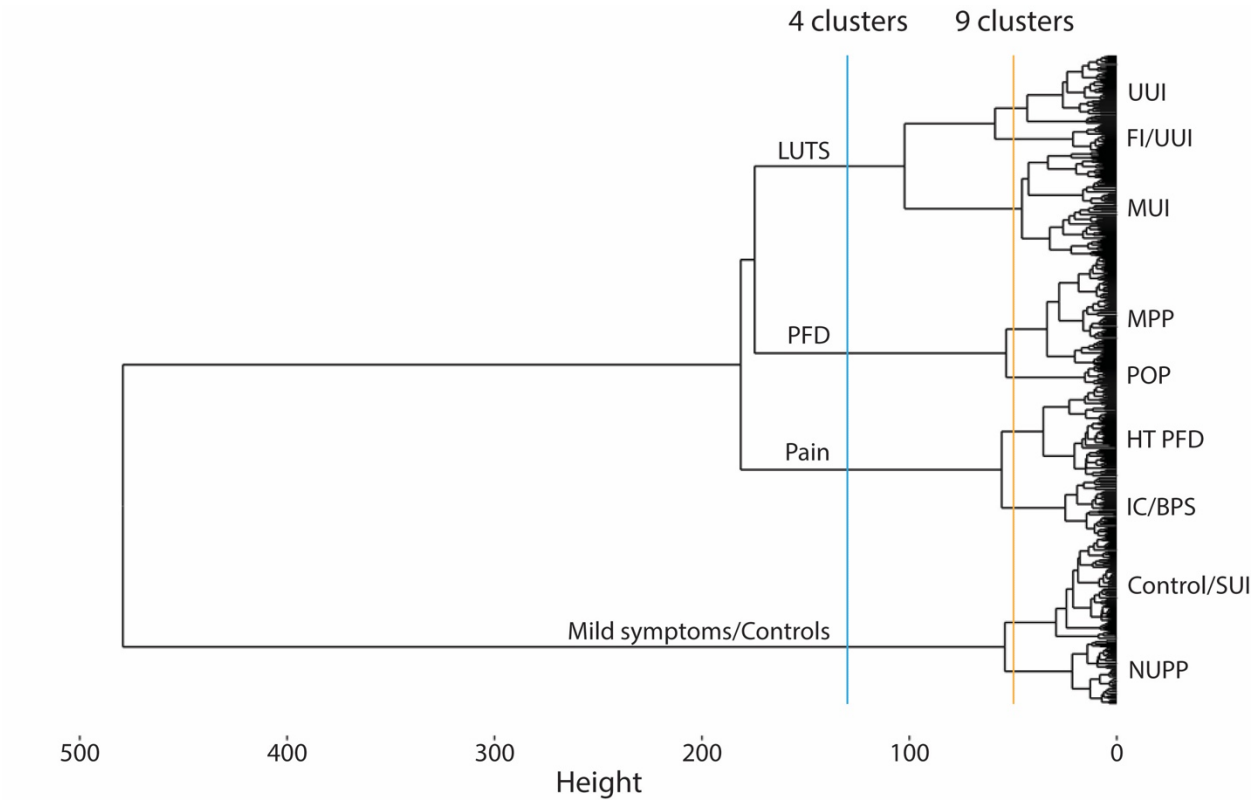


Figure 1. ***Dendrogram indicating the hierarchical relationships between clusters.*** The different lines indicate different possible cut-off points used to subclassify subjects with LUTS, with the blue line indicating the major symptomatic divisions in subjects presenting for care in a urogynecology clinic (*LUTS*: lower urinary tract symptoms; *PFD*: pelvic floor disorders; *Pain*: pelvic pain complaints; *Mild symptoms/controls*: patients with mild or minimal complaints). The yellow line indiciates the location in the dendrogram at which subjects are separated into 9 unique clusters, with each of the major categories further divided into 2-3 subclusters (UUI: urgency urinary incontinence; FI/UUI: fecal incontinence/UUI; MUI: mixed urinary incontinence;' HT PFD: high-tone pelvic floor dysfunction; POP: pelvic organ prolapse; MPP: myofascial pelvic pain; IC/BPS: interstitial cystitis/bladder pain syndrome; NUPP: non-urologic pelvic pain; Control/SUI: Asymptomatic subject/stress urinary incontinence).

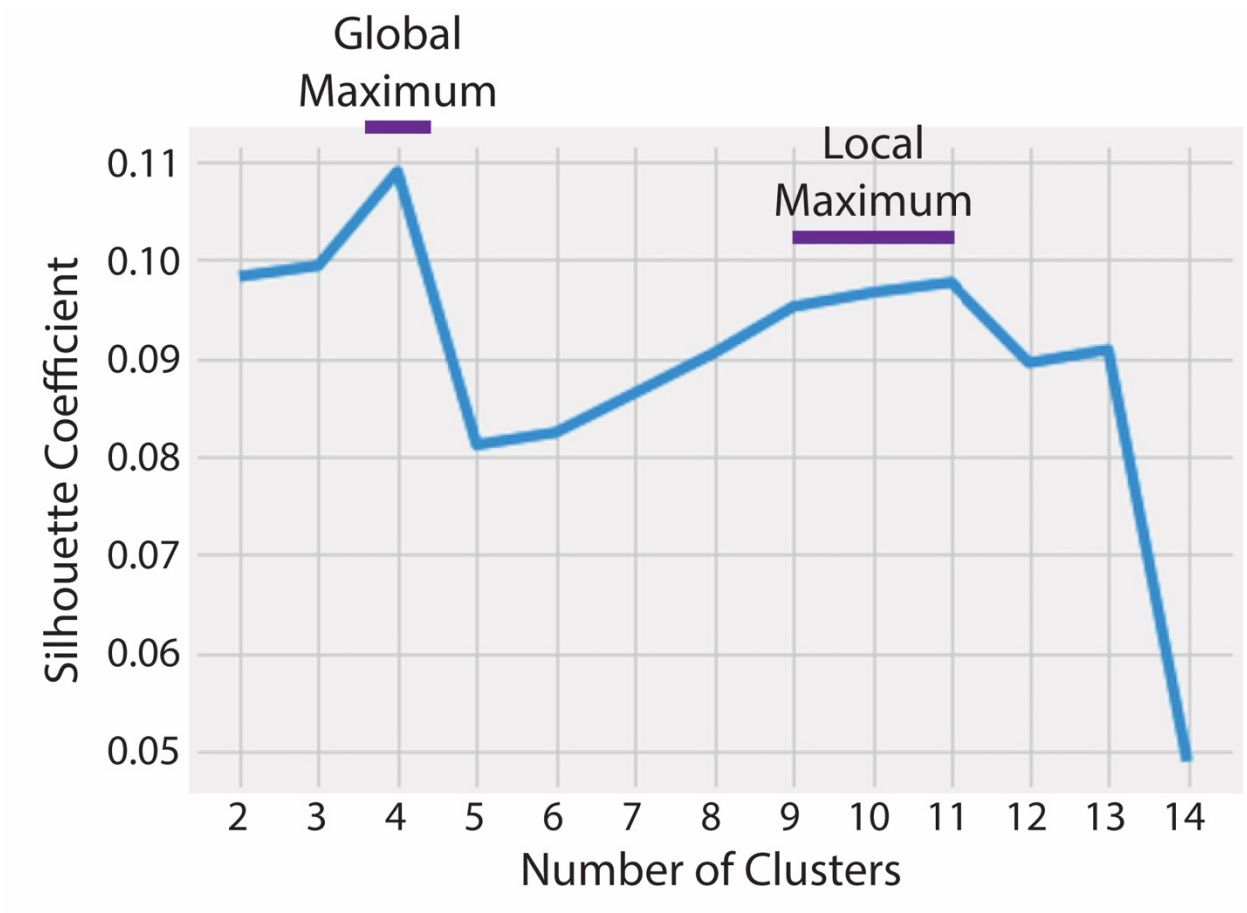


Figure 2: ***Optimal number of clusters for the hierarchical clustering by the silhouette method.*** The Silhouette score serves as a measure of the goodness of a clustering technique, encompassing both how similar each individual is to others in the same cluster (cohesion) compared to others in other distinct clusters (separation). Higher value scores indicate that each individual is well matched to their own cluster and poorly matched to neighboring clusters.

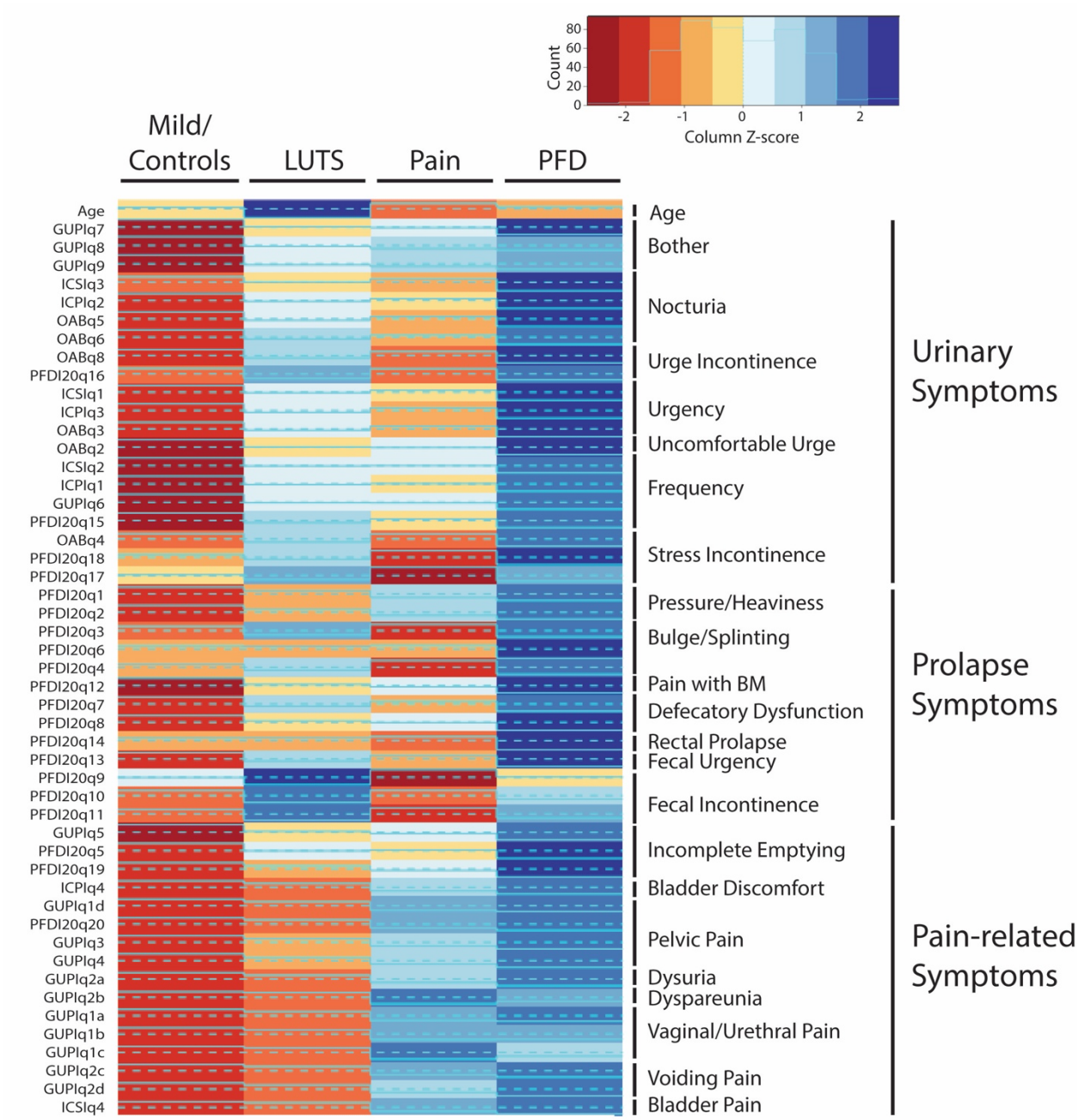


Figure 3. ***Heatmap of survey responses and patient characteristics for each of the machine learning generated clusters for the 4 major clusters.*** Specific symptomatic patterns for the 4 clusters (blue in Figure A), with deepening blues indicating scores increasingly elevated over the population mean and deeper shades of red indicating scores increasingly lower than the population mean. (*LUTS*: lower urinary tract symptoms; *PFD*: pelvic floor disorders; *Pain*: pelvic pain complaints; *Mild symptoms/controls*: patients with mild or minimal complaints).

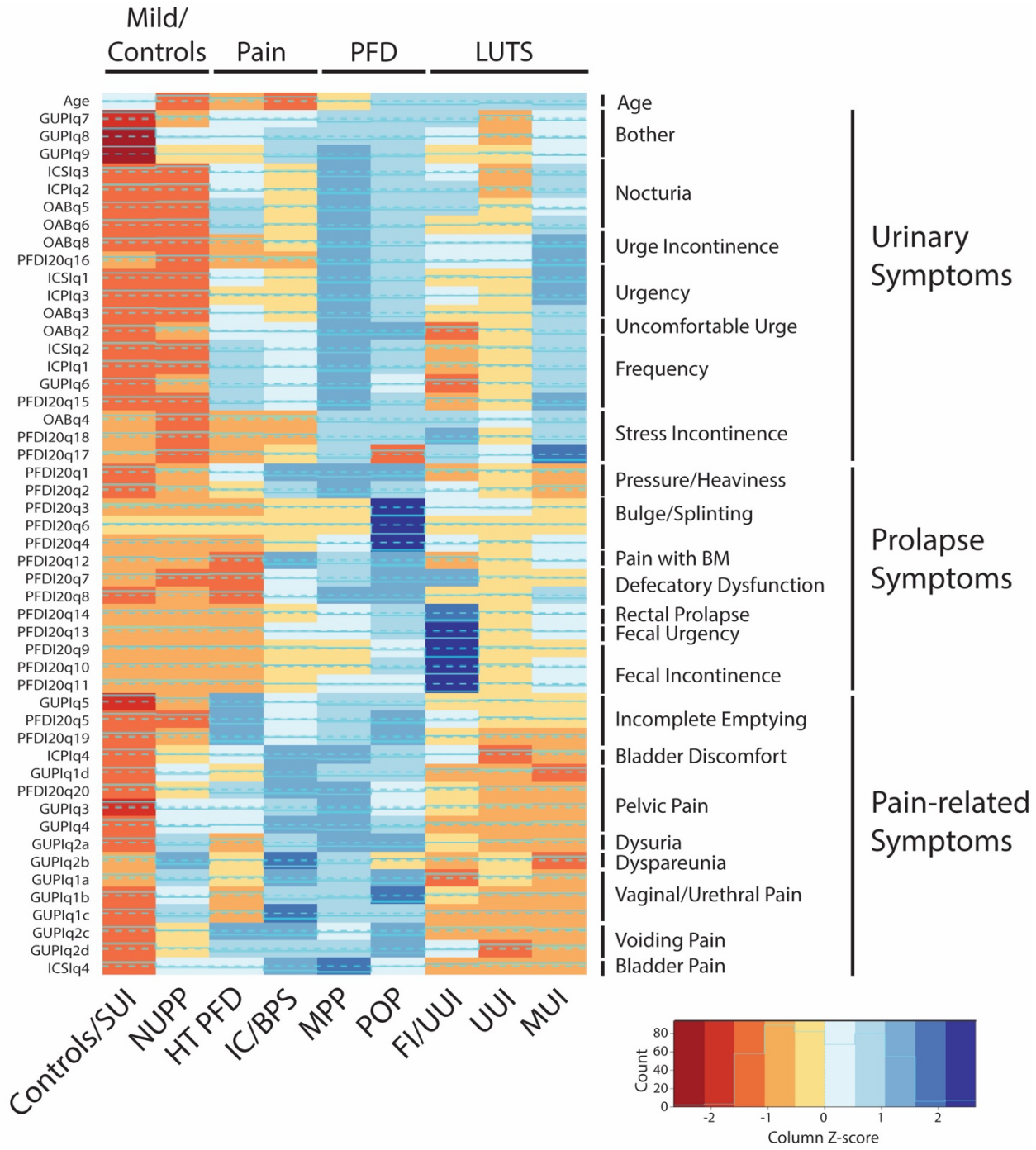


Figure 4: ***Heatmap of survey responses and patient characteristics for each of the machine learning generated clusters.*** Specific symptomatic patterns for the nine clusters shown in Figure A, with deepening blues indicating scores increasingly elevated over the population mean and deeper shades of red indicating scores increasingly lower than the population mean. (UUI: urgency urinary incontinence; FI/UUI: fecal incontinence/UUI; MUI: mixes urinary incontinence;' HT PFD: high-tone pelvic floor dysfuction; POP: pelvic organ prolapse; MPP: myofascial pelvic pain; IC/BPS: interstitial cystitis/bladder pain syndrome; NUPP: non-urologic pelvic pain; Control/SUI: Asymptomatic subject/stress urinary incontinence).

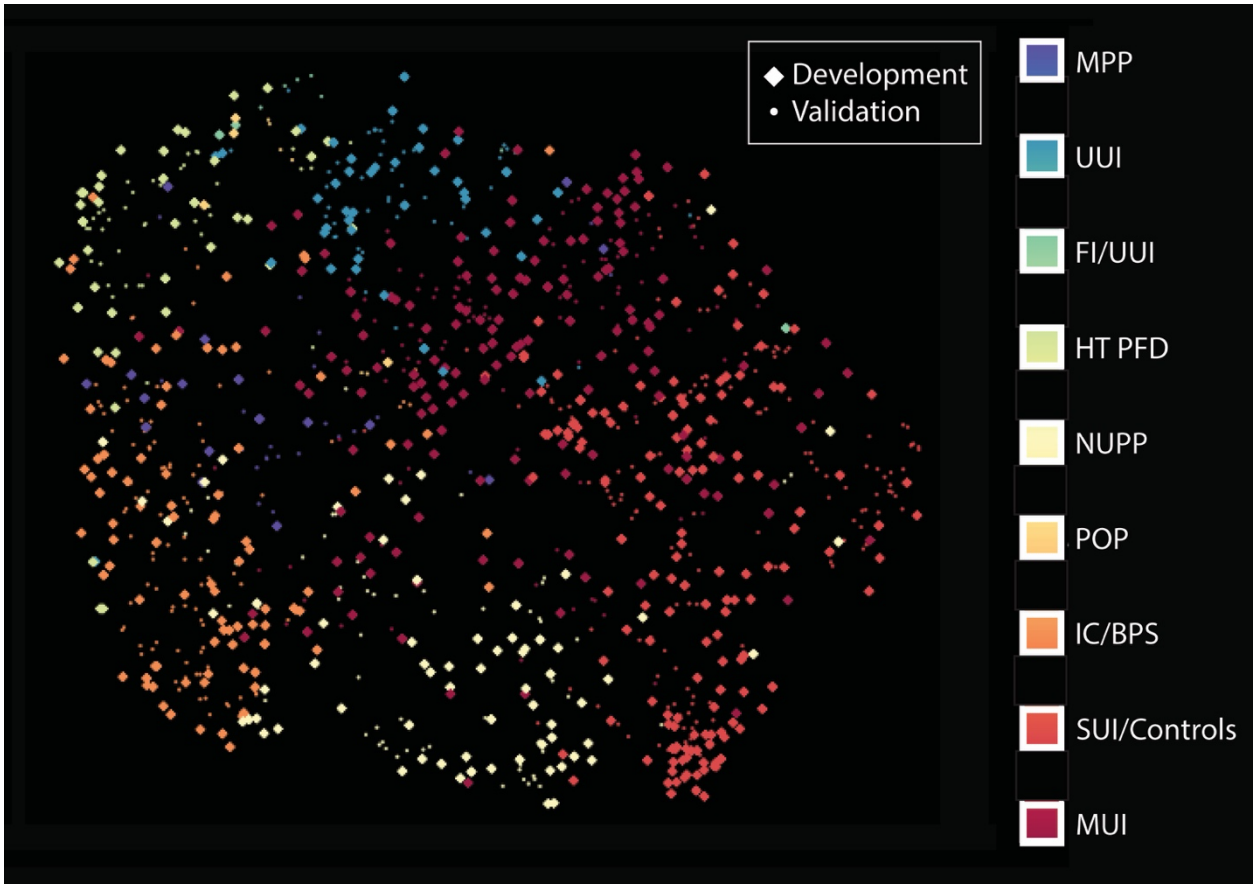


Figure 5: ***Uniform Manifold Approximation and Projection (UMAP) plot visualizing the overlap in clusters with respect to the symptomatic features between the development and validation populations.*** Good overlap was seen between the clusters in the development (larger diamonds) and validation (smaller circles) cohorts.

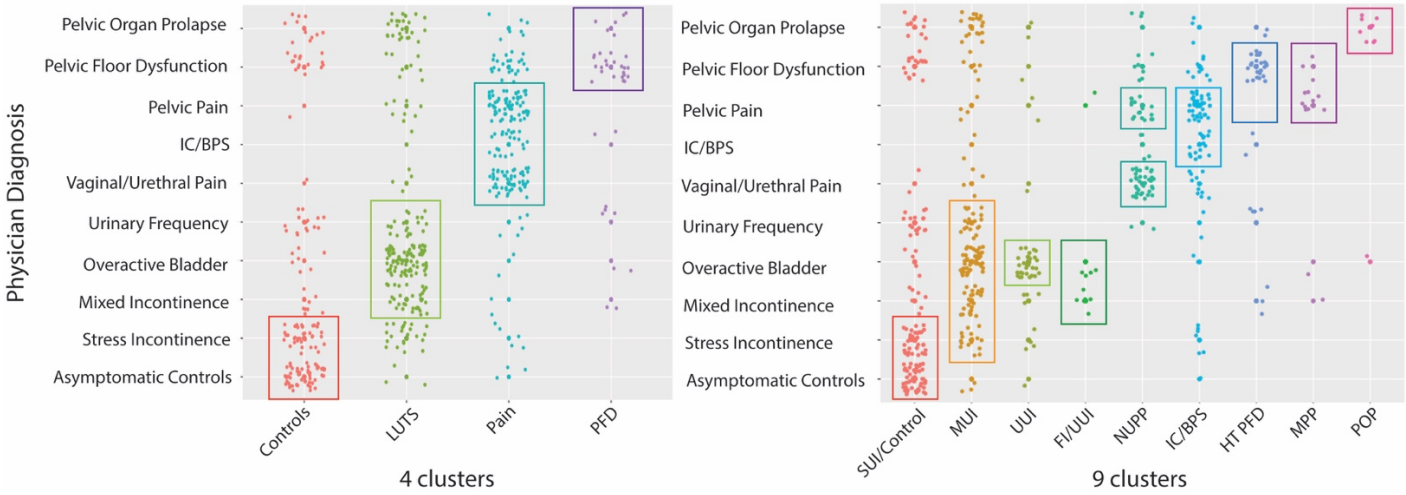
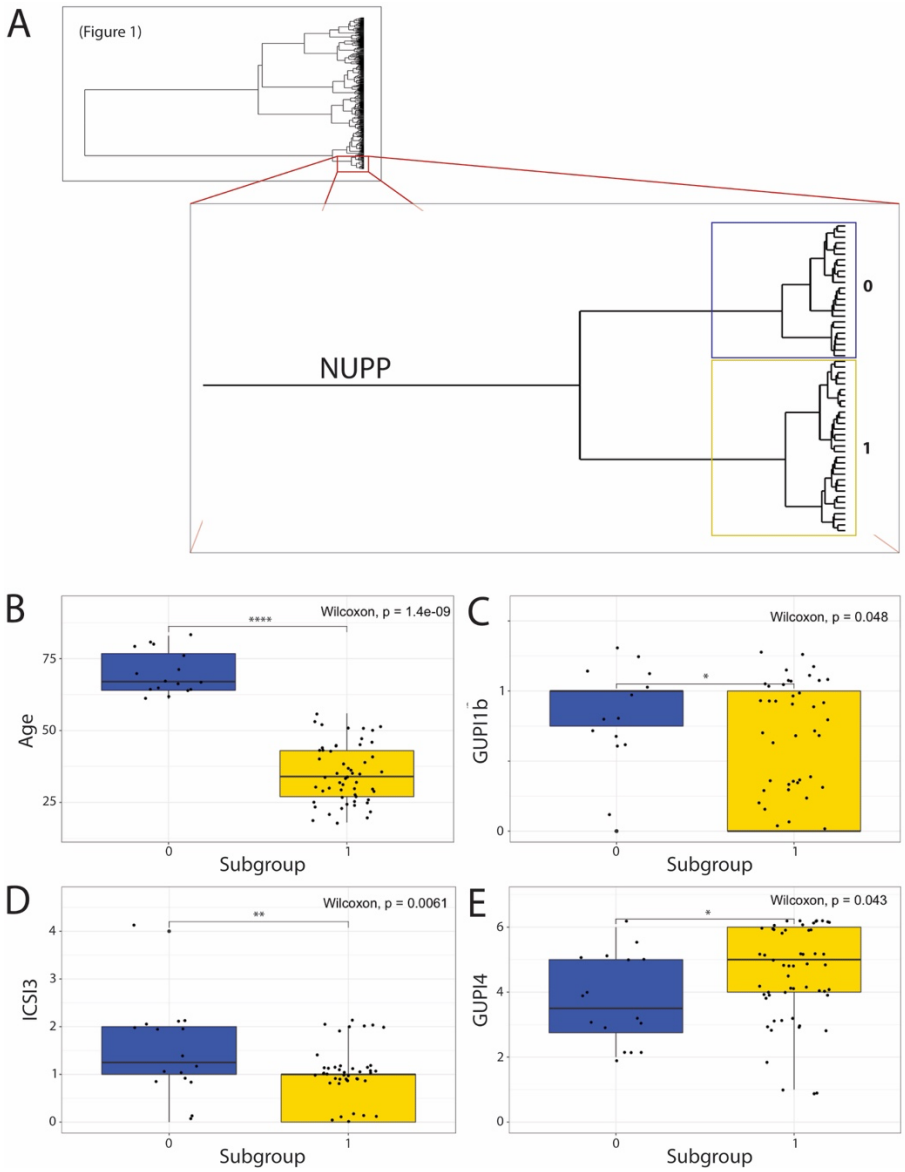


Figure 6. ***Concurrence of Physician Diagnosis with Machine Learning Clusters.*** Individual points indicate individual subjects, plotted according to the ML cluster or phenotype along the x-axis and the ICD-10 primary diagnostic code designation on the y-axis. The boxes indicate physician diagnoses determined to be in agreement with machine learning cluster phenotypes. An overall concordance greater than 70% was seen for both 4- and 9-cluster solutions.



Supplemental Figure 1. ***Symptomatic differences in the NUPP subgroups.*** A) Magnification of the region of the dendrogram representative of the NUPP cluster demonstrates the potential subdivision of patients into two subcategories, designated 0 and 1. B) While these patients tended to exhibit elevated symptom scores for dysuria, dysparaeunia, and vaginal and urethral pain, a division in this population primarily correlated with age (upper left panel). The older population exhibited more severe nocturia (ICSI q3, lower left panel) and a higher frequency of vaginal pain (fGUPI q1b, upper right panel) while the younger group had more severe pain overall (fGUPI q4, lower right panel). NUPP: non-urologic pelvic pain, *: $p<0.05$, **: $p<0.01$, ****: $p<0.001$.

Mean (SD)	Pain n = 173	PFD n = 46	LUTS n = 152	Mild Symptoms/ Controls n = 143	*p*	SMD
Age	54.38 (16.41)	57.10 (14.98)	65.37 (14.61)	57.47 (15.70)	<0.001	0.363
ICSI1	1.68 (1.44)	4.03 (1.27)	2.23 (1.64)	0.49 (0.69)	<0.001	1.537
ICSI2	3.02 (1.50)	4.43 (0.96)	3.09 (1.42)	1.50 (1.34)	<0.001	1.168
ICSI3	1.61 (1.19)	3.33 (1.19)	2.05 (1.35)	1.32 (1.19)	<0.001	0.885
ICSI4	2.04 (1.61)	2.55 (1.66)	0.44 (0.79)	0.27 (0.61)	<0.001	1.120
ICPI1	2.24 (1.25)	3.75 (0.51)	2.44 (1.22)	0.75 (0.99)	<0.001	1.627
ICPI2	1.78 (1.25)	3.46 (0.93)	2.23 (1.41)	0.88 (1.02)	<0.001	1.236
ICPI3	1.24 (1.14)	3.39 (0.69)	2.18 (1.33)	0.38 (0.69)	<0.001	1.862
ICPI4	2.67 (1.35)	3.41 (0.95)	0.90 (1.14)	0.28 (0.65)	<0.001	1.867
OABq2	3.51 (1.44)	5.35 (0.94)	3.37 (1.60)	1.54 (0.87)	<0.001	1.736
OABq3	2.37 (1.33)	5.11 (1.10)	3.40 (1.63)	1.35 (0.66)	<0.001	1.825
OABq4	1.85 (1.19)	3.96 (1.93)	3.27 (1.62)	1.81 (1.12)	<0.001	0.856
OABq5	2.85 (1.54)	4.86 (1.34)	3.61 (1.68)	1.91 (1.15)	<0.001	1.156
OABq6	3.11 (1.52)	4.90 (1.37)	3.85 (1.69)	2.06 (1.31)	<0.001	1.070
OABq8	1.73 (1.16)	4.25 (1.84)	3.16 (1.67)	1.31 (0.65)	<0.001	1.214
PFDI20 1	1.93 (1.19)	2.35 (1.26)	0.97 (1.10)	0.45 (0.54)	<0.001	1.083
PFDI20 2	1.49 (1.25)	1.90 (1.31)	0.82 (1.06)	0.38 (0.47)	<0.001	0.844
PFDI20 3	0.23 (0.45)	0.63 (0.94)	0.56 (1.06)	0.30 (0.37)	<0.001	0.327
PFDI20 4	0.41 (0.72)	0.82 (1.02)	0.69 (1.03)	0.52 (0.70)	0.003	0.270
PFDI20 5	1.37 (1.21)	2.18 (1.34)	1.44 (1.17)	0.70 (0.69)	<0.001	0.691
PFDI20 6	0.05 (0.12)	0.58 (1.18)	0.09 (0.26)	0.07 (0.17)	<0.001	0.368
PFDI20 7	0.73 (0.94)	1.28 (1.12)	1.01 (1.16)	0.50 (0.61)	<0.001	0.454
PFDI20 8	1.06 (1.04)	1.43 (1.22)	0.98 (1.05)	0.72 (0.70)	<0.001	0.364
PFDI20 9	0.09 (0.29)	0.20 (0.60)	0.31 (0.85)	0.20 (0.28)	0.007	0.208
PFDI20 10	0.30 (0.55)	0.56 (0.91)	0.66 (1.11)	0.30 (0.45)	<0.001	0.270
PFDI20 11	0.46 (0.75)	0.87 (1.00)	0.95 (1.19)	0.53 (0.55)	<0.001	0.333
PFDI20 12	0.35 (0.71)	0.51 (0.88)	0.29 (0.68)	0.13 (0.22)	0.001	0.311
PFDI20 13	0.61 (0.87)	1.09 (1.09)	0.86 (1.10)	0.49 (0.57)	<0.001	0.375
PFDI20 14	0.17 (0.51)	0.81 (1.70)	0.23 (0.61)	0.24 (0.47)	<0.001	0.281
PFDI20 15	1.93 (1.15)	2.86 (0.94)	2.32 (1.30)	1.27 (0.95)	<0.001	0.824
PFDI20 16	0.76 (0.89)	2.02 (1.43)	1.84 (1.28)	0.73 (0.68)	<0.001	0.740
PFDI20 17	0.97 (0.92)	1.55 (1.27)	1.48 (1.31)	1.29 (1.04)	<0.001	0.290
PFDI20 18	0.68 (0.78)	2.01 (1.39)	1.50 (1.36)	1.04 (0.94)	<0.001	0.651
PFDI20 19	1.02 (1.16)	1.51 (1.31)	0.67 (0.95)	0.44 (0.54)	<0.001	0.578
PFDI20 20	1.95 (1.33)	2.32 (1.19)	0.65 (0.96)	0.33 (0.39)	<0.001	1.214
fFGUPI1a	0.47 (0.48)	0.50 (0.46)	0.14 (0.32)	0.06 (0.22)	<0.001	0.730
fGUPI1b	0.52 (0.48)	0.51 (0.47)	0.12 (0.28)	0.03 (0.13)	<0.001	0.874
fGUPI1c	0.57 (0.47)	0.44 (0.44)	0.09 (0.23)	0.03 (0.13)	<0.001	0.960
fGUPI1d	0.77 (0.40)	0.85 (0.31)	0.23 (0.39)	0.06 (0.21)	<0.001	1.505
fGUPI2a	0.55 (0.48)	0.72 (0.42)	0.16 (0.35)	0.02 (0.10)	<0.001	1.189
fGUPI2b	0.52 (0.46)	0.43 (0.41)	0.12 (0.29)	0.09 (0.27)	<0.001	0.717
fGUPI2c	0.55 (0.48)	0.67 (0.44)	0.13 (0.32)	0.02 (0.12)	<0.001	1.109
fGUPI2d	0.51 (0.47)	0.68 (0.44)	0.16 (0.34)	0.06 (0.22)	<0.001	0.984
fGUPI3	3.13 (1.27)	3.94 (1.14)	1.25 (1.32)	0.44 (0.70)	<0.001	1.904
fGUPI4	5.08 (2.38)	6.57 (1.97)	2.18 (2.50)	0.56 (1.14)	<0.001	1.802
fGUPI5	2.34 (1.72)	3.59 (1.70)	1.90 (1.50)	0.55 (0.81)	<0.001	1.134
fGUPI6	2.88 (1.48)	4.11 (1.03)	2.85 (1.43)	1.22 (1.14)	<0.001	1.197
fGUPI7	1.30 (1.08)	2.26 (0.80)	1.18 (1.14)	0.18 (0.49)	<0.001	1.305
fGUPI8	2.41 (0.75)	2.67 (0.55)	2.05 (0.98)	0.55 (0.77)	<0.001	1.486

fGUPI9	4.69 (1.21)	5.34 (0.90)	4.33 (1.25)	2.01 (1.58)	<0.001	1.327

Table 1. *Patient survey scores for subjects seeking care in a urogynecology clinic by* 4-cluster machine learning diagnostic groups. Means and standard deviations are reported for each group. *LUTS*: lower urinary tract symptoms; *PFD*: pelvic floor disorders; *Pain*: pelvic pain complaints; *Mild symptoms/controls*: patients with mild or minimal complaints. *SMD*: standard mean difference.

Mean (SD)	MUI n = 100	SUI/Control n = 143	IC/BPS n = 91	POP n = 6	NUPP n = 60	MPP n = 40	FI/UUI n = 7	UUI n = 45	HT PFD n = 22	*p*	SMD
Age	65.8 (15.1)	57.5 (15.7)	50.1 (15.5)	55.8 (15.8)	59.2 (16.1)	57.3 (15.1)	57.9 (11.4)	65.5 (13.9)	59.0 (17.0)	<0.001	0.361
ICSI1	1.77 (1.30)	0.49 (0.69)	1.97 (1.48)	3.18 (0.74)	1.05 (1.17)	4.16 (1.29)	1.43 (1.90)	3.38 (1.72)	2.21 (1.42)	<0.001	1.152
ICSI2	2.65 (1.34)	1.50 (1.34)	3.50 (1.28)	4.00 (0.89)	1.95 (1.37)	4.50 (0.96)	2.22 (0.91)	4.21 (0.99)	3.99 (0.99)	<0.001	1.163
ICSI3	1.76 (1.26)	1.32 (1.19)	1.96 (1.23)	2.75 (0.42)	0.91 (0.76)	3.41 (1.24)	2.29 (1.11)	2.66 (1.40)	2.02 (1.18)	<0.001	0.864
ICSI4	0.48 (0.86)	0.27 (0.61)	2.52 (1.76)	0.89 (0.45)	1.64 (1.18)	2.80 (1.64)	0.43 (0.79)	0.33 (0.60)	1.14 (1.34)	<0.001	0.961
ICPI1	2.07 (1.14)	0.75 (0.99)	2.57 (1.01)	3.05 (0.83)	1.45 (1.29)	3.85 (0.36)	1.47 (1.39)	3.40 (0.72)	3.05 (0.90)	<0.001	1.307
ICPI2	1.86 (1.39)	0.88 (1.02)	2.18 (1.12)	3.02 (0.87)	0.88 (0.84)	3.52 (0.93)	3.00 (1.15)	2.92 (1.19)	2.59 (1.30)	<0.001	1.053
ICPI3	1.74 (1.16)	0.38 (0.69)	1.48 (1.15)	2.96 (0.95)	0.69 (0.87)	3.45 (0.64)	1.97 (1.32)	3.18 (1.15)	1.71 (1.20)	<0.001	1.338
ICPI4	0.77 (1.04)	0.28 (0.65)	3.15 (1.13)	2.44 (0.89)	2.11 (1.28)	3.55 (0.88)	1.81 (1.51)	1.05 (1.24)	2.23 (1.68)	<0.001	1.218
OABq2	2.91 (1.45)	1.54 (0.87)	3.87 (1.27)	5.00 (0.89)	2.74 (1.45)	5.40 (0.94)	2.14 (1.57)	4.58 (1.21)	4.15 (1.23)	<0.001	1.373
OABq3	2.96 (1.48)	1.35 (0.66)	2.44 (1.40)	4.33 (1.63)	1.99 (1.07)	5.22 (0.97)	3.00 (1.91)	4.44 (1.44)	3.09 (1.41)	<0.001	1.198
OABq4	2.85 (1.46)	1.81 (1.12)	1.91 (1.30)	3.83 (2.14)	1.75 (1.00)	3.98 (1.93)	4.00 (1.83)	4.08 (1.61)	1.84 (1.27)	<0.001	0.798
OABq5	3.25 (1.68)	1.91 (1.15)	3.08 (1.50)	4.33 (1.86)	1.99 (1.13)	4.94 (1.25)	4.71 (1.25)	4.25 (1.50)	4.23 (1.38)	<0.001	0.972
OABq6	3.48 (1.73)	2.06 (1.31)	3.44 (1.41)	4.33 (1.37)	2.14 (1.12)	4.99 (1.37)	3.57 (1.81)	4.71 (1.24)	4.36 (1.43)	<0.001	0.892
OABq8	2.70 (1.54)	1.31 (0.65)	1.88 (1.33)	3.33 (2.25)	1.43 (0.66)	4.38 (1.76)	2.68 (1.48)	4.26 (1.49)	1.96 (1.35)	<0.001	0.967
PFDI20 1	0.94 (1.04)	0.45 (0.54)	2.31 (1.03)	2.50 (1.64)	1.42 (1.13)	2.33 (1.21)	1.14 (1.57)	1.00 (1.19)	1.73 (1.45)	<0.001	0.741
PFDI20 2	0.82 (0.97)	0.38 (0.47)	2.01 (1.13)	1.67 (1.97)	0.85 (1.05)	1.93 (1.21)	1.57 (1.62)	0.71 (1.15)	1.05 (1.25)	<0.001	0.602
PFDI20 3	0.68 (1.20)	0.30 (0.37)	0.34 (0.55)	1.92 (1.63)	0.16 (0.26)	0.43 (0.62)	0.71 (0.95)	0.29 (0.58)	0.02 (0.10)	<0.001	0.652
PFDI20 4	0.66 (0.99)	0.52 (0.70)	0.57 (0.85)	2.17 (1.83)	0.23 (0.41)	0.62 (0.66)	1.43 (1.51)	0.65 (1.04)	0.18 (0.66)	<0.001	0.579
PFDI20 5	1.35 (1.07)	0.70 (0.69)	1.32 (0.99)	2.83 (0.98)	0.72 (0.84)	2.09 (1.37)	2.14 (1.21)	1.53 (1.34)	3.29 (0.84)	<0.001	1.070
PFDI20 6	0.07 (0.20)	0.07 (0.17)	0.06 (0.12)	3.33 (0.82)	0.05 (0.14)	0.17 (0.44)	0.14 (0.38)	0.12 (0.36)	0.00 (0.00)	<0.001	1.410
PFDI20 7	0.96 (1.11)	0.50 (0.61)	0.87 (0.92)	1.86 (1.71)	0.49 (0.73)	1.19 (1.01)	2.00 (1.73)	0.96 (1.13)	0.82 (1.33)	<0.001	0.498
PFDI20 8	0.91 (1.04)	0.72 (0.70)	1.33 (1.01)	1.67 (1.63)	0.71 (0.82)	1.40 (1.17)	1.19 (1.68)	1.09 (0.99)	0.91 (1.38)	<0.001	0.333
PFDI20 9	0.05 (0.20)	0.20 (0.28)	0.14 (0.39)	0.73 (1.61)	0.06 (0.11)	0.12 (0.18)	3.57 (0.79)	0.35 (0.68)	0.00 (0.00)	<0.001	1.560
PFDI20 10	0.25 (0.59)	0.30 (0.45)	0.37 (0.57)	1.25 (1.94)	0.28 (0.54)	0.46 (0.63)	3.71 (0.76)	1.10 (1.16)	0.11 (0.43)	<0.001	1.366
PFDI20 11	0.80 (1.09)	0.53 (0.55)	0.49 (0.73)	1.00 (1.67)	0.46 (0.72)	0.85 (0.90)	2.38 (1.57)	1.07 (1.21)	0.38 (0.97)	<0.001	0.532
PFDI20 12	0.34 (0.78)	0.13 (0.22)	0.56 (0.89)	0.77 (1.60)	0.17 (0.37)	0.47 (0.74)	0.14 (0.38)	0.18 (0.44)	0.00 (0.00)	<0.001	0.427
PFDI20 13	0.75 (1.00)	0.49 (0.57)	0.80 (0.90)	1.42 (1.86)	0.37 (0.62)	1.04 (0.96)	2.57 (1.81)	0.83 (0.98)	0.50 (1.10)	<0.001	0.569
PFDI20 14	0.14 (0.39)	0.24 (0.47)	0.16 (0.34)	0.97 (1.07)	0.12 (0.28)	0.79 (1.79)	1.43 (1.62)	0.26 (0.58)	0.36 (1.18)	<0.001	0.492
PFDI20 15	2.01 (1.22)	1.27 (0.95)	2.21 (0.84)	3.25 (0.88)	1.15 (1.05)	2.81 (0.95)	1.43 (1.81)	3.16 (0.97)	2.86 (1.32)	<0.001	0.906
PFDI20 16	1.52 (1.16)	0.73 (0.68)	0.91 (0.91)	1.58 (1.86)	0.57 (0.66)	2.08 (1.37)	1.74 (1.37)	2.58 (1.27)	0.70 (1.26)	<0.001	0.711
PFDI20 17	1.30 (1.28)	1.29 (1.04)	1.21 (0.93)	0.42 (1.02)	0.72 (0.79)	1.72 (1.23)	2.00 (1.91)	1.80 (1.21)	0.64 (0.95)	<0.001	0.567

PFDI20 18	1.20 (1.15)	1.04 (0.94)	0.79 (0.77)	2.25 (0.99)	0.61 (0.74)	1.97 (1.45)	2.89 (1.64)	1.95 (1.51)	0.45 (0.86)	<0.001	0.870
PFDI20 19	0.65 (0.90)	0.44 (0.54)	0.95 (0.92)	2.75 (0.76)	0.42 (0.62)	1.33 (1.28)	1.00 (1.41)	0.68 (1.00)	2.95 (1.17)	<0.001	1.106
PFDI20 20	0.69 (0.96)	0.33 (0.39)	2.34 (1.13)	1.92 (1.56)	1.40 (1.27)	2.38 (1.13)	1.00 (1.53)	0.50 (0.85)	1.84 (1.70)	<0.001	0.836
fGUPI1a	0.21 (0.37)	0.06 (0.22)	0.52 (0.48)	0.61 (0.44)	0.51 (0.48)	0.48 (0.47)	0.03 (0.07)	0.01 (0.06)	0.16 (0.35)	<0.001	0.810
fGUPI1b	0.15 (0.31)	0.03 (0.13)	0.58 (0.47)	0.61 (0.44)	0.55 (0.48)	0.49 (0.48)	0.17 (0.37)	0.04 (0.16)	0.16 (0.35)	<0.001	0.774
fGUPI1c	0.11 (0.26)	0.03 (0.13)	0.72 (0.43)	0.44 (0.46)	0.50 (0.47)	0.44 (0.45)	0.07 (0.11)	0.04 (0.16)	0.14 (0.30)	<0.001	0.857
fGUPI1d	0.28 (0.42)	0.06 (0.21)	0.92 (0.25)	0.66 (0.41)	0.70 (0.42)	0.87 (0.29)	0.14 (0.38)	0.12 (0.30)	0.30 (0.45)	<0.001	1.207
fGUPI2a	0.18 (0.36)	0.02 (0.10)	0.71 (0.44)	0.57 (0.50)	0.43 (0.48)	0.75 (0.41)	0.29 (0.49)	0.10 (0.29)	0.18 (0.35)	<0.001	0.842
fGUPI2b	0.15 (0.31)	0.09 (0.27)	0.65 (0.44)	0.32 (0.40)	0.43 (0.46)	0.45 (0.41)	0.14 (0.38)	0.06 (0.21)	0.24 (0.38)	<0.001	0.619
fGUPI2c	0.08 (0.24)	0.02 (0.12)	0.76 (0.42)	0.55 (0.50)	0.25 (0.41)	0.68 (0.43)	0.14 (0.38)	0.23 (0.42)	0.52 (0.49)	<0.001	0.854
fGUPI2d	0.14 (0.31)	0.06 (0.22)	0.64 (0.46)	0.73 (0.44)	0.37 (0.45)	0.67 (0.44)	0.43 (0.53)	0.17 (0.37)	0.37 (0.45)	<0.001	0.731
fGUPI3	1.34 (1.33)	0.44 (0.70)	3.39 (1.16)	2.79 (0.98)	2.91 (1.25)	4.11 (1.07)	1.49 (1.80)	1.01 (1.21)	2.69 (1.53)	<0.001	1.250
fGUPI4	2.47 (2.67)	0.56 (1.14)	5.80 (1.93)	4.52 (2.07)	4.56 (2.54)	6.88 (1.79)	2.14 (2.85)	1.55 (1.93)	3.48 (2.55)	<0.001	1.220
fGUPI5	1.95 (1.46)	0.55 (0.81)	2.64 (1.66)	3.41 (1.57)	1.33 (1.36)	3.62 (1.74)	2.00 (1.63)	1.77 (1.59)	3.82 (1.26)	<0.001	0.950
fGUPI6	2.40 (1.25)	1.22 (1.14)	3.48 (1.18)	3.01 (1.21)	1.62 (1.23)	4.28 (0.91)	1.71 (1.38)	4.04 (1.04)	3.79 (0.94)	<0.001	1.208
fGUPI7	0.83 (0.91)	0.18 (0.49)	1.45 (1.05)	2.02 (1.25)	1.11 (1.13)	2.29 (0.72)	2.14 (1.46)	1.82 (1.19)	1.24 (0.97)	<0.001	0.860
fGUPI8	1.88 (0.94)	0.55 (0.77)	2.49 (0.67)	2.55 (0.50)	2.25 (0.84)	2.69 (0.56)	2.43 (1.13)	2.35 (0.98)	2.47 (0.77)	<0.001	0.793
fGUPI9	3.99 (1.20)	2.01 (1.58)	4.91 (1.07)	4.93 (1.11)	4.35 (1.43)	5.40 (0.86)	4.43 (1.62)	5.06 (0.96)	4.68 (0.88)	<0.001	0.812

Table 2. ***Group symptom means and differences for 9 symptom subclusters of subjects seeking care in a urogynecology clinic***. Means and standard deviations are reported for each group. MUI: mixed urinary incontinence; Control/SUI: Asymptomatic subject/stress urinary incontinence; IC/BPS: interstitial cystitis/bladder pain syndrome; POP: pelvic organ prolapse; NUPP: non-urologic pelvic pain; MPP: myofascial pelvic pain; FI/UUI: fecal incontinence/urgency urinary incontinence; UUI: urgency urinary incontinence; HT PFD: high-tone pelvic floor dysfunction. SMD: Standard mean difference.

Mean (SD)	MUI n = 143	SUI/Control n = 151	IC/BPS n = 92	POP n = 9	NUPP n =71	MPP n = 35	FI/UUI n = 11	UUI n = 45	HT PFD n = 14	*p*	SMD
Age	59.34 (17.17)	54.97 (15.93)	45.39 (16.48)	61.73 (19.15)	42.77 (17.56)	54.31 (20.26)	61.84 (10.80)	62.51 (20.39)	48.93 (15.69)	<0.001	0.521
ICSI1	2.04 (1.41)	0.55 (0.81)	2.21 (1.66)	3.34 (0.69)	0.63 (0.78)	4.10 (1.32)	1.73 (1.95)	3.78 (1.43)	2.43 (1.45)	<0.001	1.267
ICSI2	2.74 (1.47)	1.32 (1.33)	3.61 (1.47)	3.89 (0.78)	1.90 (1.40)	4.69 (0.68)	2.23 (0.88)	4.21 (0.94)	4.21 (0.97)	<0.001	1.288
ICSI3	1.55 (1.22)	1.30 (1.18)	1.96 (1.21)	2.78 (0.36)	0.96 (0.78)	3.19 (1.33)	2.36 (1.12)	2.80 (1.15)	2.32 (1.10)	<0.001	0.874
ICSI4	0.38 (0.66)	0.23 (0.49)	1.66 (1.25)	0.93 (0.36)	1.06 (1.07)	1.90 (1.40)	0.55 (0.82)	0.38 (0.83)	1.11 (1.24)	<0.001	0.736
ICPI1	2.20 (1.26)	0.82 (0.99)	2.78 (0.98)	3.03 (0.83)	1.16 (1.08)	3.89 (0.40)	1.48 (1.50)	3.44 (0.78)	3.00 (1.04)	<0.001	1.316
ICPI2	1.76 (1.42)	1.02 (1.17)	2.27 (1.26)	3.01 (0.85)	0.92 (0.98)	3.37 (1.03)	3.00 (1.18)	3.17 (1.04)	2.79 (1.48)	<0.001	0.951
ICPI3	1.94 (1.33)	0.42 (0.78)	1.74 (1.28)	2.98 (0.90)	0.42 (0.69)	3.57 (0.61)	2.07 (1.39)	3.36 (0.91)	1.71 (1.14)	<0.001	1.456
ICPI4	0.70 (1.04)	0.37 (0.87)	3.02 (1.21)	2.52 (0.92)	1.80 (1.42)	3.40 (1.06)	1.88 (1.54)	0.93 (1.39)	2.07 (1.73)	<0.001	1.057
OABq2	3.01 (1.46)	1.77 (1.11)	3.98 (1.43)	5.11 (0.93)	2.69 (1.58)	5.17 (1.15)	2.27 (1.62)	4.77 (1.18)	3.86 (1.35)	<0.001	1.173
OABq3	3.19 (1.52)	1.47 (0.85)	3.02 (1.66)	4.67 (1.50)	1.76 (1.02)	5.17 (1.10)	3.00 (1.90)	4.51 (1.50)	3.36 (1.55)	<0.001	1.134
OABq4	3.08 (1.53)	1.97 (1.22)	2.20 (1.44)	4.22 (1.99)	1.54 (0.95)	4.06 (1.98)	4.18 (1.72)	4.21 (1.68)	2.14 (1.75)	<0.001	0.830
OABq5	3.14 (1.66)	2.04 (1.33)	3.27 (1.70)	4.67 (1.66)	1.87 (1.26)	5.11 (1.13)	4.82 (1.25)	4.42 (1.74)	4.64 (1.28)	<0.001	1.013
OABq6	3.37 (1.64)	2.38 (1.37)	3.66 (1.62)	4.56 (1.24)	2.23 (1.27)	5.20 (1.13)	3.82 (1.66)	4.91 (1.41)	4.64 (1.34)	<0.001	0.922
OABq8	2.81 (1.55)	1.41 (0.77)	2.29 (1.60)	3.89 (2.15)	1.31 (0.77)	4.20 (1.97)	2.88 (1.64)	4.07 (1.57)	1.64 (1.39)	<0.001	0.913
PFDI20 1	1.20 (1.38)	0.32 (0.77)	2.60 (1.26)	2.67 (1.58)	1.06 (1.27)	2.91 (1.27)	1.09 (1.64)	1.09 (1.40)	1.79 (1.37)	<0.001	0.816
PFDI20 2	0.97 (1.29)	0.14 (0.52)	2.01 (1.46)	2.00 (2.00)	0.68 (1.13)	2.37 (1.50)	1.64 (1.69)	0.49 (1.06)	1.29 (1.27)	<0.001	0.718
PFDI20 3	0.81 (1.40)	0.18 (0.63)	0.56 (1.11)	2.11 (1.34)	0.23 (0.66)	0.51 (1.07)	0.73 (0.90)	0.58 (1.18)	0.21 (0.80)	<0.001	0.546
PFDI20 4	0.59 (1.15)	0.31 (0.85)	0.74 (1.26)	2.33 (1.87)	0.18 (0.48)	1.03 (1.29)	1.09 (1.38)	0.84 (1.35)	0.29 (0.83)	<0.001	0.555
PFDI20 5	1.47 (1.43)	0.40 (0.80)	2.07 (1.35)	3.00 (1.00)	0.73 (1.09)	2.63 (1.54)	2.27 (1.10)	1.74 (1.54)	3.29 (0.83)	<0.001	1.052
PFDI20 6	0.24 (0.65)	0.11 (0.54)	0.22 (0.55)	3.00 (1.32)	0.12 (0.36)	0.26 (0.78)	0.09 (0.30)	0.31 (0.82)	0.00 (0.00)	<0.001	0.816
PFDI20 7	0.86 (1.21)	0.48 (1.05)	1.16 (1.23)	1.79 (1.71)	0.44 (0.99)	1.69 (1.37)	1.82 (1.72)	1.07 (1.37)	0.36 (0.74)	<0.001	0.545
PFDI20 8	0.86 (1.13)	0.43 (0.93)	1.33 (1.27)	1.67 (1.66)	0.52 (0.94)	1.69 (1.62)	1.03 (1.55)	1.51 (1.29)	0.50 (0.85)	<0.001	0.485
PFDI20 9	0.21 (0.67)	0.08 (0.42)	0.22 (0.68)	1.04 (1.71)	0.07 (0.24)	0.40 (0.91)	3.55 (0.82)	0.64 (1.09)	0.00 (0.00)	<0.001	1.309
PFDI20 10	0.41 (0.95)	0.22 (0.69)	0.64 (1.15)	1.67 (1.98)	0.15 (0.50)	0.63 (1.09)	3.64 (0.81)	1.38 (1.34)	0.00 (0.00)	<0.001	1.298
PFDI20 11	0.86 (1.25)	0.35 (0.77)	0.97 (1.38)	1.11 (1.76)	0.41 (0.92)	1.11 (1.39)	2.61 (1.33)	1.33 (1.60)	0.46 (1.18)	<0.001	0.582
PFDI20 12	0.34 (0.79)	0.13 (0.56)	0.90 (1.31)	1.07 (1.70)	0.25 (0.67)	0.83 (1.34)	0.09 (0.30)	0.56 (1.08)	0.00 (0.00)	<0.001	0.488
PFDI20 13	0.91 (1.26)	0.36 (0.91)	1.28 (1.42)	1.78 (1.91)	0.39 (0.80)	1.40 (1.44)	3.00 (1.55)	1.31 (1.47)	0.57 (1.16)	<0.001	0.691
PFDI20 14	0.29 (0.79)	0.13 (0.56)	0.28 (0.77)	0.93 (1.11)	0.14 (0.46)	0.74 (1.20)	1.64 (1.63)	0.71 (1.22)	0.00 (0.00)	<0.001	0.591
PFDI20 15	1.89 (1.38)	0.77 (1.13)	2.35 (1.21)	3.33 (0.83)	0.75 (1.19)	3.63 (0.91)	1.27 (1.79)	3.56 (0.87)	3.00 (1.24)	<0.001	1.199
PFDI20 16	1.69 (1.40)	0.49 (0.94)	0.73 (1.12)	2.11 (1.75)	0.17 (0.48)	2.57 (1.67)	1.93 (1.43)	3.00 (1.28)	0.43 (0.85)	<0.001	1.014
PFDI20 17	1.38 (1.54)	0.87 (1.27)	1.11 (1.45)	0.56 (1.10)	0.48 (1.01)	1.89 (1.57)	1.82 (1.78)	2.33 (1.43)	0.79 (1.12)	<0.001	0.567

PFDI20 18	1.35 (1.44)	0.64 (1.03)	0.88 (1.31)	2.11 (0.89)	0.38 (0.88)	2.34 (1.66)	2.65 (1.73)	2.22 (1.57)	0.71 (1.27)	<0.001	0.784
PFDI20 19	0.69 (1.13)	0.28 (0.69)	1.41 (1.44)	2.78 (0.79)	0.32 (0.87)	2.26 (1.75)	1.09 (1.51)	0.71 (1.29)	2.71 (1.38)	<0.001	1.000
PFDI20 20	0.69 (1.18)	0.13 (0.52)	2.55 (1.24)	2.00 (1.56)	1.25 (1.42)	2.83 (1.27)	1.09 (1.58)	0.73 (1.29)	2.46 (1.52)	<0.001	0.925
fGUPI1a	0.18 (0.39)	0.06 (0.24)	0.54 (0.50)	0.63 (0.46)	0.42 (0.50)	0.46 (0.51)	0.02 (0.06)	0.09 (0.29)	0.21 (0.43)	<0.001	0.702
fGUPI1b	0.13 (0.33)	0.04 (0.20)	0.55 (0.50)	0.74 (0.40)	0.41 (0.50)	0.46 (0.51)	0.20 (0.40)	0.13 (0.34)	0.07 (0.27)	<0.001	0.744
fGUPI1c	0.12 (0.32)	0.02 (0.14)	0.74 (0.44)	0.51 (0.48)	0.54 (0.50)	0.51 (0.51)	0.13 (0.30)	0.11 (0.32)	0.14 (0.36)	<0.001	0.762
fGUPI1d	0.24 (0.43)	0.03 (0.16)	0.89 (0.31)	0.77 (0.37)	0.61 (0.49)	0.77 (0.43)	0.18 (0.40)	0.07 (0.25)	0.36 (0.50)	<0.001	1.105
fGUPI2a	0.18 (0.39)	0.01 (0.11)	0.67 (0.47)	0.71 (0.45)	0.56 (0.50)	0.69 (0.47)	0.36 (0.50)	0.09 (0.29)	0.07 (0.27)	<0.001	0.890
fGUPI2b	0.23 (0.42)	0.10 (0.30)	0.67 (0.47)	0.21 (0.35)	0.54 (0.50)	0.49 (0.51)	0.18 (0.40)	0.09 (0.29)	0.29 (0.47)	<0.001	0.569
fGUPI2c	0.08 (0.28)	0.03 (0.16)	0.61 (0.49)	0.59 (0.50)	0.28 (0.45)	0.40 (0.50)	0.09 (0.30)	0.18 (0.39)	0.57 (0.51)	<0.001	0.690
fGUPI2d	0.10 (0.31)	0.02 (0.14)	0.56 (0.50)	0.71 (0.45)	0.34 (0.48)	0.57 (0.50)	0.45 (0.52)	0.20 (0.40)	0.57 (0.51)	<0.001	0.694
fGUPI3	1.41 (1.24)	0.43 (0.70)	3.22 (1.10)	2.75 (1.09)	2.76 (1.03)	3.81 (1.20)	1.68 (1.90)	1.14 (1.45)	2.64 (1.39)	<0.001	1.108
fGUPI4	2.44 (2.51)	0.66 (1.50)	5.52 (1.97)	4.46 (2.24)	4.19 (1.85)	6.26 (2.20)	2.36 (2.91)	1.78 (2.08)	3.50 (2.47)	<0.001	1.036
fGUPI5	1.89 (1.38)	0.48 (0.81)	2.64 (1.67)	3.44 (1.65)	1.45 (1.46)	3.34 (1.81)	2.18 (1.60)	2.29 (1.81)	3.93 (1.14)	<0.001	0.922
fGUPI6	2.54 (1.34)	1.14 (1.19)	3.53 (1.21)	3.06 (1.35)	1.82 (1.40)	4.40 (0.77)	1.64 (1.43)	3.99 (1.27)	4.07 (1.00)	<0.001	1.180
fGUPI7	1.04 (0.98)	0.41 (0.73)	1.76 (0.96)	2.01 (1.31)	1.18 (1.17)	2.17 (0.89)	2.18 (1.40)	1.89 (0.98)	1.71 (1.14)	<0.001	0.671
fGUPI8	1.93 (0.87)	0.97 (0.95)	2.58 (0.63)	2.59 (0.49)	2.35 (0.81)	2.71 (0.52)	2.27 (1.19)	2.44 (0.76)	2.29 (0.73)	<0.001	0.672
fGUPI9	3.93 (1.24)	2.47 (1.68)	4.91 (1.18)	4.84 (1.17)	4.32 (1.42)	5.43 (0.81)	4.27 (1.74)	4.64 (1.16)	4.32 (1.44)	<0.001	0.676

Supplemental Table 1. ***Group symptom means and differences for 9 symptom subclusters of the validation cohort.*** Means and standard deviations are reported for each group in the validation cohort. MUI: mixed urinary incontinence; Control/SUI: Asymptomatic subject/stress urinary incontinence; IC/BPS: interstitial cystitis/bladder pain syndrome; POP: pelvic organ prolapse; NUPP: non-urologic pelvic pain; MPP: myofascial pelvic pain; FI/UUI: fecal incontinence/urgency urinary incontinence; UUI: urgency urinary incontinence; HT PFD: high-tone pelvic floor dysfunction. SMD: Standard mean difference.

ICD 10 code	Corresponding Diagnosis	4-cluster assignment	9-cluster assignment
N81	Pelvic Organ Prolapse	PFD	POP
M99.05, N94.2	Pelvic Floor Dysfunction	PFD	MPP, HT PFD
R10.2, N94.1	Pelvic Pain	Pain	IC/BPS, MPP, NUPP
N30.1	IC/BPS	Pain	IC/BPS
R10.2, N94.81	Vaginal Pain, Urethral Pain	Pain	NUPP
R35.0	Urinary Frequency	LUTS	MUI
N32.81	Overactive Bladder	LUTS	UUI, MUI, FI/UUI
N39.46	Mixed Incontinence	LUTS	MUI
N39.3	Stress Incontinence	Mild Symptoms/ Controls	SUI/Controls
N/A	Asymptomatic Controls	Mild Symptoms/ Controls	SUI/Controls

Supplemental Table 2. ***Ten most common diagnoses for patients in the validation cohort.*** Each diagnosis, determined from the associated ICD 10 codes linked to the visit at which these patients provided the patient-reported questionnaire data, was assigned to an ML symptom cluster assignment to determine approximate accuracy of the various models.

Mean (SD)	Cystocele/MUI n = 18	MUI/SUI n = 82	*p* (vs. MUI)	High-grade POP n = 6	*p* (vs. POP)
Age	**60.95 (13.03)**	66.88 (15.37)	0.963	55.76 (15.78)	0.017
ICSI1	**1.17 (0.92)**	1.91 (1.34)	0.002	3.18 (0.74)	0.001
ICSI2	**2.61 (1.54)**	2.66 (1.30)	0.090	4.00 (0.89)	0.026
ICSI3	**1.39 (0.98)**	1.84 (1.30)	0.813	2.75 (0.42)	0.002
ICSI4	**0.22 (0.43)**	0.54 (0.92)	0.788	0.89 (0.45)	0.407
ICPI1	**1.72 (1.13)**	2.14 (1.14)	0.092	3.05 (0.83)	0.021
ICPI2	**1.61 (1.33)**	1.91 (1.40)	0.667	3.02 (0.87)	0.021
ICPI3	**1.17 (1.15)**	1.87 (1.13)	0.040	2.96 (0.95)	0.011
ICPI4	**1.00 (1.03)**	0.71 (1.04)	0.049	2.44 (0.89)	0.055
OABq2	**2.33 (1.53)**	3.04 (1.41)	0.021	5.00 (0.89)	0.001
OABq3	**2.06 (1.35)**	3.16 (1.44)	0.0002	4.33 (1.63)	0.001
OABq4	**2.26 (1.44)**	2.98 (1.44)	0.008	3.83 (2.14)	0.002
OABq5	**2.94 (1.47)**	3.32 (1.72)	0.858	4.33 (1.86)	0.004
OABq6	**3.28 (1.67)**	3.52 (1.75)	0.844	4.33 (1.37)	0.019
OABq8	**1.72 (1.23)**	2.92 (1.52)	0.0001	3.33 (2.25)	0.0005
PFDI20 1	**1.94 (1.30)**	0.72 (0.83)	<0.0001	2.50 (1.64)	0.051
PFDI20 2	**1.61 (1.46)**	0.65 (0.72)	<0.0001	1.67 (1.97)	0.073
PFDI20 3	**3.00 (1.03)**	0.17 (0.30)	<0.0001	1.92 (1.63)	0.716
PFDI20 4	**0.78 (1.35)**	0.63 (0.89)	0.955	2.17 (1.83)	0.003
PFDI20 5	**1.33 (1.46)**	1.36 (0.98)	0.379	2.83 (0.98)	0.011
PFDI20 6	**0.11 (0.32)**	0.06 (0.16)	0.100	3.33 (0.82)	0.009
PFDI20 7	**0.44 (1.04)**	1.07 (1.10)	0.333	1.86 (1.71)	0.016
PFDI20 8	**1.06 (1.26)**	0.88 (0.99)	0.717	1.67 (1.63)	0.065
PFDI20 9	**0.00 (0.00)**	0.07 (0.22)	0.424	0.73 (1.61)	0.0002
PFDI20 10	**0.13 (0.47)**	0.27 (0.61)	0.373	1.25 (1.94)	0.0005
PFDI20 11	**0.67 (1.24)**	0.83 (1.06)	0.825	1.00 (1.67)	0.133
PFDI20 12	**0.11 (0.47)**	0.40 (0.83)	0.664	0.77 (1.60)	0.003
PFDI20 13	**0.53 (1.04)**	0.80 (0.99)	0.181	1.42 (1.86)	0.006
PFDI20 14	**0.01 (0.03)**	0.16 (0.43)	0.788	0.97 (1.07)	0.010
PFDI20 15	**1.56 (1.46)**	2.11 (1.14)	0.061	3.25 (0.88)	0.005
PFDI20 16	**1.22 (1.35)**	1.59 (1.11)	0.013	1.58 (1.86)	0.002
PFDI20 17	**1.44 (1.54)**	1.27 (1.23)	0.994	0.42 (1.02)	0.404
PFDI20 18	**0.61 (0.98)**	1.33 (1.15)	0.201	2.25 (0.99)	0.018
PFDI20 19	**0.78 (1.06)**	0.62 (0.86)	0.052	2.75 (0.76)	0.003
PFDI20 20	**0.78 (1.17)**	0.67 (0.92)	0.089	1.92 (1.56)	0.042
fGUPI1a	**0.23 (0.42)**	0.20 (0.36)	0.197	0.61 (0.44)	0.050
fGUPI1b	**0.18 (0.38)**	0.14 (0.30)	0.048	0.61 (0.44)	0.003
fGUPI1c	**0.01 (0.05)**	0.14 (0.28)	0.616	0.44 (0.46)	0.002
fGUPI1d	**0.28 (0.46)**	0.28 (0.41)	0.410	0.66 (0.41)	0.006
fGUPI2a	**0.17 (0.38)**	0.18 (0.35)	0.954	0.57 (0.50)	0.013
fGUPI2b	**0.18 (0.38)**	0.14 (0.30)	0.741	0.32 (0.40)	0.310
fGUPI2c	**0.11 (0.32)**	0.08 (0.22)	0.405	0.55 (0.50)	0.016
fGUPI2d	**0.06 (0.24)**	0.16 (0.33)	0.407	0.73 (0.44)	0.004
fGUPI3	**1.42 (1.28)**	1.33 (1.35)	0.466	2.79 (0.98)	0.054
fGUPI4	**2.53 (2.75)**	2.45 (2.67)	0.843	4.52 (2.07)	0.019
fGUPI5	**1.44 (1.34)**	2.06 (1.47)	0.604	3.41 (1.57)	0.057
fGUPI6	**2.06 (1.26)**	2.47 (1.24)	0.085	3.01 (1.21)	0.073
fGUPI7	**0.83 (0.92)**	0.83 (0.91)	0.481	2.02 (1.25)	0.106
fGUPI8	**1.89 (1.08)**	1.88 (0.91)	0.381	2.55 (0.50)	0.094
fGUPI9	**3.87 (1.04)**	4.01 (1.24)	0.699	4.93 (1.11)	0.073

Supplemental Table 3. ***Group symptom means and pairwise comparisons for MUI alone and Cystocele/MUI subgroups***. Means and standard deviations are reported for each group. *P* values were determined by Wilcoxon signed-rank test. MUI: mixed urinary incontinence; Cystocele/MUI: MUI subjects with anterior prolapse. The high-grade POP group is shown for comparison.

REFERENCES

1. Coyne KS, Sexton CC, Thompson CL, Milsom I, Irwin D, Kopp ZS, et al. The prevalence of lower urinary tract symptoms (LUTS) in the USA, the UK and Sweden: results from the Epidemiology of LUTS (EpiLUTS) study. BJU Int. 2009;104(3):352-60.
2. Lee UJ, Feinstein L, Ward JB, Kirkali Z, Martinez-Miller EE, Matlaga BR, et al. Prevalence of Urinary Incontinence among a Nationally Representative Sample of Women, 2005-2016: Findings from the Urologic Diseases in America Project. J Urol. 2021;205(6):1718-24.
3. Coyne KS, Sexton CC, Irwin DE, Kopp ZS, Kelleher CJ, Milsom I. The impact of overactive bladder, incontinence and other lower urinary tract symptoms on quality of life, work productivity, sexuality and emotional well-being in men and women: results from the EPIC study. BJU Int. 2008;101(11):1388-95.
4. Coyne KS, Wein A, Nicholson S, Kvasz M, Chen CI, Milsom I. Comorbidities and personal burden of urgency urinary incontinence: a systematic review. Int J Clin Pract. 2013;67(10):1015-33.
5. Coyne KS, Wein A, Nicholson S, Kvasz M, Chen CI, Milsom I. Economic burden of urgency urinary incontinence in the United States: a systematic review. Journal of managed care pharmacy : JMCP. 2014;20(2):130-40.
6. Berry SH, Bogart LM, Pham C, Liu K, Nyberg L, Stoto M, et al. Development, validation and testing of an epidemiological case definition of interstitial cystitis/painful bladder syndrome. J Urol. 2010;183(5):1848-52.
7. Benner JS, Nichol MB, Rovner ES, Jumadilova Z, Alvir J, Hussein M, et al. Patient-reported reasons for discontinuing overactive bladder medication. BJU Int. 2010;105(9):1276-82.
8. NIH. Meeeting on Measurement of Urinary Symptoms [Available from: http://www.niddk.nih.gov/news/events-calendar/Pages/meeting-on-measurement-of-urinary-symptoms-momus.aspx.
9. Ackerman AL, Lai HH, Parameshwar PS, Eilber KS, Anger JT. Symptomatic Overlap in Overactive Bladder and Interstitial Cystitis/Painful Bladder Syndrome - development of a new algorithm. BJU Int. 2018.
10. Lai HH, Vetter J, Jain S, Gereau RWt, Andriole GL. The overlap and distinction of self-reported symptoms between interstitial cystitis/bladder pain syndrome and overactive bladder: a questionnaire based analysis. J Urol. 2014;192(6):1679-85.
11. Liedl B, Goeschen K, Sutherland SE, Roovers JP, Yassouridis A. Can surgical reconstruction of vaginal and ligamentous laxity cure overactive bladder symptoms in women with pelvic organ prolapse? BJU Int. 2019;123(3):493-510.
12. Karjalainen PK, Tolppanen AM, Mattsson NK, Wihersaari OAE, Jalkanen JT, Nieminen K. Pelvic organ prolapse surgery and overactive bladder symptoms-a population-based cohort (FINPOP). Int Urogynecol J. 2022;33(1):95-105.
13. Clemens JQ, Erickson DR, Varela NP, Lai HH. Diagnosis and Treatment of Interstitial Cystitis/Bladder Pain Syndrome. J Urol. 2022;208(1):34-42.
14. Lightner DJ, Gomelsky A, Souter L, Vasavada SP. Diagnosis and Treatment of Overactive Bladder (Non-Neurogenic) in Adults: AUA/SUFU Guideline Amendment 2019. J Urol. 2019;202(3):558-63.
15. Blaivas JG, Li ESW, Dayan L, Edeson ME, Mathew J, O'Boyle AL, et al. Overactive bladder phenotypes: development and preliminary data. Can J Urol. 2021;28(3):10699-704.
16. Herschorn S. Overactive bladder: Symptom complex or separate entity? Canadian Urological Association journal = Journal de l'Association des urologues du Canada. 2011;5(5 Suppl 2):S152-4.

17. Griebling TL. Geriatric Urology. In: Solomon DH, LoCicero J, Rosenthal RA, American Geriatrics Society, Foundation JAH, editors. New Frontiers in Geriatrics Research: An Agenda for Surgical and Related Medical Specialties. New York: American Geriatrics Society; 2004. p. 269-302.
18. Adamson AS, Welch HG. Machine Learning and the Cancer-Diagnosis Problem - No Gold Standard. N Engl J Med. 2019;381(24):2285-7.
19. Awan SE, Sohel F, Sanfilippo FM, Bennamoun M, Dwivedi G. Machine learning in heart failure: ready for prime time. Curr Opin Cardiol. 2018;33(2):190-5.
20. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine Learning for Medical Imaging. Radiographics. 2017;37(2):505-15.
21. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. N Engl J Med. 2016;375(13):1216-9.
22. Suarez-Ibarrola R, Hein S, Reis G, Gratzke C, Miernik A. Current and future applications of machine and deep learning in urology: a review of the literature on urolithiasis, renal cell carcinoma, and bladder and prostate cancer. World J Urol. 2020;38(10):2329-47.
23. Checcucci E, Autorino R, Cacciamani GE, Amparore D, De Cillis S, Piana A, et al. Artificial intelligence and neural networks in urology: current clinical applications. Minerva Urol Nefrol. 2020;72(1):49-57.
24. Choo MS, Uhmn S, Kim JK, Han JH, Kim DH, Kim J, et al. A Prediction Model Using Machine Learning Algorithm for Assessing Stone-Free Status after Single Session Shock Wave Lithotripsy to Treat Ureteral Stones. J Urol. 2018;200(6):1371-7.
25. Goldenberg SL, Nir G, Salcudean SE. A new era: artificial intelligence and machine learning in prostate cancer. Nat Rev Urol. 2019;16(7):391-403.
26. Hasnain Z, Mason J, Gill K, Miranda G, Gill IS, Kuhn P, et al. Machine learning models for predicting post-cystectomy recurrence and survival in bladder cancer patients. PloS one. 2019;14(2):e0210976.
27. Kocak B, Yardimci AH, Bektas CT, Turkcanoglu MH, Erdim C, Yucetas U, et al. Textural differences between renal cell carcinoma subtypes: Machine learning-based quantitative computed tomography texture analysis with independent external validation. Eur J Radiol. 2018;107:149-57.
28. Ozkan IA, Koklu M, Sert IU. Diagnosis of urinary tract infection based on artificial intelligence methods. Comput Methods Programs Biomed. 2018;166:51-9.
29. Sheyn D, Ju M, Zhang S, Anyaeche C, Hijaz A, Mangel J, et al. Development and Validation of a Machine Learning Algorithm for Predicting Response to Anticholinergic Medications for Overactive Bladder Syndrome. Obstet Gynecol. 2019;134(5):946-57.
30. Karmonik C, Boone T, Khavari R. Data-Driven Machine-Learning Quantifies Differences in the Voiding Initiation Network in Neurogenic Voiding Dysfunction in Women With Multiple Sclerosis. Int Neurourol J. 2019;23(3):195-204.
31. Clemens J, Calhoun E, Litwin M, McNaughton-Collins M, Kusek J, Crowley E, et al. Validation of a modified National Institutes of Health chronic prostatitis symptom index to assess genitourinary pain in both men and women. Urology. 2009;74:983-7.e3.
32. Coyne K, Revicki D, Hunt T, Corey R, Stewart W, Bentkover J, et al. Psychometric validation of an overactive bladder symptom and health-related quality of life questionnaire: The OAB-q. . Qual Life Res. 2002;11:563-74.
33. Barber MD, Walters MD, Bump RC. Short forms of two condition-specific quality-of-life questionnaires for women with pelvic floor disorders (PFDI-20 and PFIQ-7). Am J Obstet Gynecol. 2005;193(1):103-13.
34. O'Leary MP, Sant GR, Fowler FJ, Whitmore KE, Spolarich-Kroll J. The interstitial cystitis symptom index and problem index. Urology. 1997;49(5):58-63.
35. Ward JHJ. Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association. 1963;58(301):236-44.

36. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics. 1987;20:53-65.
37. Grinsztajn L, Oyallon E, Varoquaux G. Why do tree-based models still outperform deep learning on tabular data? arXiv [Internet]. 2022.
38. Brodersen KH, Ong CS, Stephan KE, Buhmann JM. The Balanced Accuracy and Its Posterior Distribution. 2010 20th International Conference on Pattern Recognition; 23-26 Aug. 20102010. p. 3121-4.
39. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv [Internet]. 2018.
40. Hall SA, Cinar A, Link CL, Kopp ZS, Roehrborn CG, Kaplan SA, et al. Do urological symptoms cluster among women? Results from the Boston Area Community Health Survey. BJU Int. 2008;101(10):1257-66.
41. Andreev VP, Helmuth ME, Liu G, Smith AR, Merion RM, Yang CC, et al. Subtyping of common complex diseases and disorders by integrating heterogeneous data. Identifying clusters among women with lower urinary tract symptoms in the LURN study. PloS one. 2022;17(6):e0268547.
42. Andreev VP, Liu G, Yang CC, Smith AR, Helmuth ME, Wiseman JB, et al. Symptom Based Clustering of Women in the LURN Observational Cohort Study. J Urol. 2018;200(6):1323-31.
43. Coyne KS, Matza LS, Kopp ZS, Thompson C, Henry D, Irwin DE, et al. Examining lower urinary tract symptom constellations using cluster analysis. BJU Int. 2008;101(10):1267-73.
44. Ueda T, Hanno PM, Saito R, Meijlink JM, Yoshimura N. Current Understanding and Future Perspectives of Interstitial Cystitis/Bladder Pain Syndrome. Int Neurourol J. 2021;25(2):99-110.
45. Lai HH, Jemielita T, Sutcliffe S, Bradley CS, Naliboff B, Williams DA, et al. Characterization of Whole Body Pain in Urological Chronic Pelvic Pain Syndrome at Baseline: A MAPP Research Network Study. J Urol. 2017;198(3):622-31.
46. Gross J, Vetter JM, Lai HH. Clustering of patients with overactive bladder syndrome. BMC urology. 2021;21(1):41.
47. Mwesigwa PJ, Jackson NJ, Caron AT, Kanji F, Ackerman JE, Webb JR, et al. Unsupervised Machine Learning Approaches Reveal Distinct Phenotypes of Perceived Bladder Pain: A Pilot Study. Front Pain Res (Lausanne). 2021;2.
48. Chow D, Rodriguez LV. Epidemiology and prevalence of pelvic organ prolapse. Current opinion in urology. 2013;23(4):293-8.
49. Bradley CS, Zimmerman MB, Qi Y, Nygaard IE. Natural history of pelvic organ prolapse in postmenopausal women. Obstet Gynecol. 2007;109(4):848-54.
50. Handa VL, Garrett E, Hendrix S, Gold E, Robbins J. Progression and remission of pelvic organ prolapse: a longitudinal study of menopausal women. Am J Obstet Gynecol. 2004;190(1):27-32.
51. Bharucha AE, Dunivan G, Goode PS, Lukacz ES, Markland AD, Matthews CA, et al. Epidemiology, pathophysiology, and classification of fecal incontinence: state of the science summary for the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) workshop. The American journal of gastroenterology. 2015;110(1):127-36.
52. Bharucha AE, Knowles CH, Mack I, Malcolm A, Oblizajek N, Rao S, et al. Faecal incontinence in adults. Nat Rev Dis Primers. 2022;8(1):53.
53. Quaghebeur J, Petros P, Wyndaele JJ, De Wachter S. Pelvic-floor function, dysfunction, and treatment. Eur J Obstet Gynecol Reprod Biol. 2021;265:143-9.
54. Sinha S. Dysfunctional voiding: A review of the terminology, presentation, evaluation and management in children and adults. Indian J Urol. 2011;27(4):437-47.

55. Voorham-van der Zalm PJ, Lycklama ANGAB, Elzevier HW, Putter H, Pelger RCM. "Diagnostic investigation of the pelvic floor": a helpful tool in the approach in patients with complaints of micturition, defecation, and/or sexual dysfunction. J Sex Med. 2008;5(4):864-71.
56. Meister MR, Sutcliffe S, Badu A, Ghetti C, Lowder JL. Pelvic floor myofascial pain severity and pelvic floor disorder symptom bother: is there a correlation? Am J Obstet Gynecol. 2019;221(3):235 e1- e15.
57. Ross V, Detterman C, Hallisey A. Myofascial Pelvic Pain: An Overlooked and Treatable Cause of Chronic Pelvic Pain. J Midwifery Womens Health. 2021;66(2):148-60.
58. Westbay LC, Adams W, Kistner M, Brincat C, Bresler L, Yang LC, et al. Clinical Outcomes of a Multidisciplinary Female Chronic Pelvic Pain Program. Female Pelvic Med Reconstr Surg. 2021.
59. Chronic Pelvic Pain: ACOG Practice Bulletin, Number 218. Obstet Gynecol. 2020;135(3):e98-e109.
60. Ahangari A. Prevalence of chronic pelvic pain among women: an updated review. Pain Physician. 2014;17(2):E141-7.
61. Latthe P, Latthe M, Say L, Gulmezoglu M, Khan KS. WHO systematic review of prevalence of chronic pelvic pain: a neglected reproductive health morbidity. BMC Public Health. 2006;6:177.
62. Clemens JQ, Stephens-Shields A, Naliboff BD, Lai HH, Rodriguez L, Krieger JN, et al. Correlates of Health Care Seeking Activities in Patients with Urological Chronic Pelvic Pain Syndromes: Findings from the MAPP Cohort. J Urol. 2018;200(1):136-40.
63. Reiter RC. A profile of women with chronic pelvic pain. Clin Obstet Gynecol. 1990;33(1):130-6.
64. Kaplan JA, Simianu VV. Pelvic Floor Nonrelaxation: Approach to Evaluation and Treatment. Clin Colon Rectal Surg. 2021;34(1):49-55.
65. Kotarinos RK. Pelvic floor physical therapy in urogynecologic disorders. Curr Womens Health Rep. 2003;3(4):334-9.
66. FitzGerald MP, Kotarinos R. Rehabilitation of the short pelvic floor. I: Background and patient evaluation. International urogynecology journal and pelvic floor dysfunction. 2003;14(4):261-8.
67. Louis-Charles K, Biggie K, Wolfinbarger A, Wilcox B, Kienstra CM. Pelvic Floor Dysfunction in the Female Athlete. Curr Sports Med Rep. 2019;18(2):49-52.
68. Butrick CW. Pathophysiology of pelvic floor hypertonic disorders. Obstet Gynecol Clin North Am. 2009;36(3):699-705.
69. Faubion SS, Shuster LT, Bharucha AE. Recognition and management of nonrelaxing pelvic floor dysfunction. Mayo Clinic proceedings. 2012;87(2):187-93.
70. Ackerman AL, Torosis M, Jackson NJ, Routh JC, Lowder JL. Myofascial Frequency Syndrome: the red herring of pelvic conditions. 2022.
71. Wolff BJ, Joyce CJ, Brincat CA, Mueller ER, Fitzgerald CM. Pelvic floor myofascial pain in patients with symptoms of urinary tract infection. Int J Gynaecol Obstet. 2019;145(2):205-11.
72. Schussler-Fiorenza Rose SM, Gangnon RE, Chewning B, Wald A. Increasing Discussion Rates of Incontinence in Primary Care: A Randomized Controlled Trial. J Womens Health (Larchmt). 2015;24(11):940-9.
73. Malani P, Kullgren J, Solway E. National Poll on Healthy Aging (NPHA). Ann Arbor, MI: University of Michigan; 2019.