HHS Public Access
Author manuscript
J Urol. Author manuscript; available in PMC 2019 December 01.

Published in final edited form as:
J Urol. 2018 December ; 200(6): 1323–1331. doi:10.1016/j.juro.2018.06.068.

SYMPTOM-BASED CLUSTERING OF WOMEN IN THE SYMPTOMS OF LOWER URINARY TRACT DYSFUNCTION RESEARCH NETWORK (LURN) OBSERVATIONAL COHORT STUDY

Victor P. Andreev, PhD[1], **Gang Liu, PhD**[1], **Claire C. Yang, MD**[2], **Abigail R Smith, PhD**[1], **Margaret E. Helmuth, MA**[1], **Jonathan B. Wiseman, MS**[1], **Robert M. Merion, MD FACS**[1], **Kevin P. Weinfurt, PhD**[3], **Anne P. Cameron, MD**[4], **H. Henry Lai, MD**[5], **David Cella, PhD**[6], **Brenda W. Gillespie, PhD**[4], **Brian T. Helfand, MD PhD**[7], **James W. Griffith, PhD**[6], **John O. L. DeLancey, MD**[4], **Matthew O. Fraser, PhD**[3], **J. Quentin Clemens, MD FACS MSCI**[4], **Ziya Kirkali, MD**[8], and **LURN Study Group**

[1]Arbor Research Collaborative for Health, Ann Arbor, MI

[2]University of Washington, Seattle, WA

[3]Duke University Medical Center

[4]University of Michigan, Ann Arbor, MI

[5]Washington University School of Medicine, St. Louis, MO

[6]Northwestern University Feinberg School of Medicine, Chicago, IL

[7]NorthShore University Health System, Glenview, IL

[8]National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD

Abstract

Purpose—Women with lower urinary tract symptoms (LUTS) are often diagnosed based on a pre-defined symptom complex, or on a predominant symptom. There are many limitations to this paradigm, as often patients present with multiple urinary symptoms that do not perfectly fit the pre-established diagnoses. We utilized cluster analysis to identify novel symptom-based subtypes of women with LUTS.

Materials and Methods—Baseline urinary symptom questionnaire data were analyzed from 545 care-seeking female participants enrolled in the LURN Observational Cohort Study. Symptoms were measured using the LUTS Tool and the AUA Symptom Index and analyzed using a probability-based consensus clustering algorithm.

Corresponding Author: Victor P. Andreev PhD, Arbor Research Collaborative for Health 340 East Huron Street, Suite 300, Ann Arbor, MI 48104 USA, Tel: (734) 665-4108, Direct: (734) 369-9386, Fax: (734) 665-2103, Victor.Andreev@ArborResearch.orgwww.ArborResearch.org. Gang.liu@arborresearch.org, cyang@u.washington.edu, abby.smith@arborresearch.org, Margaret.helmuth@arborresearch.org, Jon.wiseman@arborresearch.org, Robert.merion@arborresearch.org, Kevin.Weinfurt@duke.edu, annepell@med.umich.edu, laih@wudosis.wustl.edu, d-cella@northwestern.edu, bgillesp@umich.edu, bhelfand@northshore.org, J-Griffith@northwestern.edu, delancey@med.umich.edu, Matthew.Fraser@duke.edu, qclemens@med.umich.edu, ziya.kirkali@nih.gov.

Results—Four clusters were identified. Women in cluster F1 (n=138) do not report incontinence, but experience post-void dribbling, frequency, and voiding symptoms. Women in cluster F2 (n=80) report urgency incontinence, as well as urgency and frequency, and very minimal voiding symptoms or stress incontinence. Cluster F3 (n=244) includes women who report all types of incontinence, urgency, frequency, and very mild voiding symptoms. Women in cluster F4 (n=83) report all LUTS at uniformly high levels. All but two of 44 LUTS Tool and 8 AUA symptom questions were significantly ($p<0.05$) different between at least two clusters, and all clusters contained at least one member from each conventional group (continent, stress incontinence, urgency incontinence, mixed incontinence, and other incontinence).

Conclusions—Women seeking care for LUTS cluster into four distinct symptom groups that differ from conventional clinical diagnostic groups. Further validation is needed to determine whether management improves with this new classification.

Keywords

symptom-based subtypes; overactive bladder; nocturia; stress incontinence; urgency incontinence; consensus clustering

Introduction

The current paradigm for managing patients with lower urinary tract symptoms (LUTS) is to assign a diagnosis based on a pre-defined symptom complex, such as overactive bladder (OAB), or based on a single predominant symptom, such as nocturia or stress urinary incontinence (SUI). Treatments are then administered based on these diagnoses. But there are limitations to this paradigm, as patients frequently present with other urinary symptoms in addition to those being treated, and these combinations of symptoms may be relevant to treatment selection. Diagnosis and treatment based solely on patients' predominant symptoms may be unsatisfactory, as it disregards other presenting symptoms.

Mechanistic studies reveal that a functional impairment to a specific organ in the urinary tract may cause more than a single symptom. For example, a weak urethra is associated with both SUI and urge urinary incontinence (UUI)[1,2], resulting in the common complaint of mixed urinary incontinence (MUI). This raises the question of how current diagnostic paradigms correspond with biological changes of continence system and how symptoms occur in women seeking treatment. Moreover, mixed symptoms other than MUI may be common, presenting a complex combination of urinary incontinence, voiding, and storage symptoms that are especially difficult to treat.

Previous investigations into novel LUTS subtypes include the EPIC and BACH studies, which aimed to define subtypes based on a relatively small number of self-reported symptom data in community dwelling cohorts.[3,4] Another study of a large cohort of treatment seeking women focused only on the patients with overactive bladder and sought to identify groups of highly correlated symptoms.[5] Little work has been done to group patients from a care-seeking population.

The Symptoms of Lower Urinary Tract Dysfunction Research Network (LURN) is a multi-center study funded by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).[6] One goal of LURN was to improve the treatment of patients with LUTS by improving our understanding of the types of patients presenting with LUTS. As part of the LURN Observational Cohort Study[7], we captured urinary symptoms from a large group of symptomatic women presenting for care. These data were subjected to cluster analysis, an approach based on a distribution of multiple symptoms, rather than the identification of predominant symptoms. By clustering groups of individuals with similar patterns of data, in this case LUTS, our main objective was to identify clusters that may better represent LUTS subtypes.

Materials and Methods

Data

Data were obtained from the LURN Observational Cohort Study[4], which included 545 women presenting with LUTS at six US tertiary care centers and has been described previously.[6,7] Baseline data collection (i.e., prior to treatment by a LURN physician) included demographic information, clinical exam findings, and several questionnaires. The LUTS Tool[8,9] and the American Urological Association Symptom Index (AUA-SI)[10] were the primary source of LUTS data. The LUTS Tool has 44 items, including questions on severity and bother for each symptom. The AUA-SI has seven items; the quality of life question later added was also included.[11] Participants also completed patient reported outcome (PRO) questionnaires related to bowel function (PROMIS gastrointestinal constipation, diarrhea, and bowel incontinence subsets)[12], psychological health (PROMIS Depression and Anxiety Short Forms[13], Perceived Stress Scale[14], PROMIS Sleep Disturbance Short Form[15]), urologic pain (Genitourinary Pain Index [GUPI])[16], and pelvic floor function (Pelvic Floor Distress Inventory [PFDI])[17]. The decision to cluster patients only on urologic symptoms from the LUTS Tool and AUA-SI was made prior to the analysis.

Methods

Responses to the LUTS Tool and AUA-SI were occasionally missing (up to 10% per question), therefore, multiple imputation was performed. Five imputed data sets were created using sequential regression techniques implemented in IVEware version 2.0[18,19].

To avoid clustering predominantly by the overall severity of LUTS, we normalized the data by the participant’s overall severity (see details in the Supplemental Material). We also accounted for correlation among the items, partly due to redundancy in the questions (Supplemental Figure 1). Clustering results can be skewed by including variables reflecting redundant information, therefore, we weighted items so that the highest weight was attributed to the least correlated question (i.e., the question with smallest average correlation with all other questions) and lowest weight to the most correlated question as defined by equations (S1) and (S2) in the Supplemental Material.

Clustering was performed using a resampling based consensus clustering method introduced by Monti et al[20]. The details of this procedure are presented in the Supplemental Material. The final step resulted in the generation of the consensus matrix (Figure 1), where each element of the matrix can be interpreted as the probability, for each pair of participants, of belonging to the same cluster. The above process was performed for each of the five imputed data sets and probabilities were averaged across data sets, allowing seamless integration of the multiple imputations. Clusters were then formed by using the above probabilities to group participants. Resulting clusters were examined using “quality of clustering” criteria that compares between-cluster differences to within-cluster differences[21–26]. Quality of clustering was optimal for the number of clusters equal to four (Supplemental Table 1). Differences across clusters in demographics, LUTS, and other PROs (listed in the Data section above) were examined by using one-way ANOVA, chi-square tests, and multiple linear regression as described in the Supplemental Material.

Results

Four distinct clusters of individuals were identified by analyzing responses of 545 female participants to the LUTS Tool and AUA-SI questionnaires. (We call these clusters: F1-F4 in order to distinguish them from the clusters of male patients (M1-M4) defined in our ongoing study.) The consensus matrix (Figure 1) shows high contrast between on-diagonal (yellow) and off-diagonal (blue) blocks, demonstrating the unambiguous results of the clustering. Demographic characteristics were not different across the clusters, with the exceptions of obesity (BMI>30kg/m^2) and vaginal births (Table 1). Obesity presented at significantly higher frequencies within cluster F3 (55% compared to 30% and 40% in clusters F1 and F2, p<0.001 and p=0.02, respectively) and cluster F4 (47% compared to 30% in cluster F1, p=0.01, see Table 1 for overall p-values), whereas vaginal parity was significantly higher in clusters 3 and 4 (76% relative to 63% and 68% in clusters 1 and 2). In contrast, urinary symptoms were quite different across the clusters. No cluster could be characterized by a single symptom, but rather by a combination of symptoms with various levels of severity, shown using radar plots in Figure 2. Women in cluster F1 (n=138) did not report incontinence, but had post-void dribbling, frequency, and voiding symptoms. Women in cluster F2 (n=80) reported urgency incontinence, as well as urgency and frequency, and very minimal voiding symptoms or stress incontinence. Cluster F3 (n=244), which is the most populated cluster in our cohort, included women who reported all types of incontinence (including SUI), urgency, frequency, and very mild voiding symptoms. Women in cluster F4 (n=83) reported all LUTS at uniformly high levels.

Multiple urinary symptoms are present at significantly different levels across the clusters. The LUTS Tool severity questions tended to be rated significantly higher by participants in cluster F4, with the exception of daytime frequency, which was similar to other clusters, and urgency which was rated lower compared to clusters F2 and F3. Post-micturition, pain, and incontinence symptoms were rated significantly higher in cluster F3 compared to cluster F2 (Figure 3).

Non-urologic PROs and other urologic PROs not used for clustering were also present at significantly different severity levels across clusters (Figure 4). Comparison of each pair of clusters demonstrates at least four and up to twelve significantly different scores (upper triangle of the matrix). Cluster F4 tended to have higher (more severe) scores for all PROs and these were significantly different from at least one other cluster for all measures except the Perceived Stress Scale. Clusters F1, F2, and F3 mostly different on the GUPI and PFDI-20 and associated subscales, however, cluster F3 also had more severe diarrhea compared to cluster F1 and more severe sleep disturbance compared to cluster F2.

The presence of multiple significantly different symptoms across the clusters illustrates that our clusters meet the concise definition of clustering given by Liao[27] as: “The goal of clustering is to identify structure in an unlabeled data set by objectively organizing data into homogeneous groups where the within-group-object dissimilarity is minimized and the between-group-object dissimilarity is maximized.”

Using conventional incontinence groups based on a subset of incontinence questions from LUTS Tool, we classified each participant as continent, stress urinary incontinent, urgency urinary incontinent, mixed urinary incontinent, and other urinary incontinent. At least one patient from each of the conventional incontinence groups was represented in each of the four clusters (Figure 5). The “quality of clustering” criteria was higher for the new clusters than the conventional groups (Supplemental Table 2). In addition, significant differences in non-urologic PROs were more common between clusters than between conventional groups (Supplemental Table 3).

Discussion

We identified four clusters of treatment-seeking female LUTS patients. Pairwise comparison of clusters demonstrated clear distinctions in LUTS distribution between the four clusters, as well as multiple significant differences of non-urological symptoms. Participants in cluster F1 would previously have been labelled OAB dry because there was minimal incontinence, but that classification ignored the voiding and post-micturition symptoms. In addition, urgency is the defining diagnostic criterion for OAB, which is not central to this cluster. Participants in cluster F2 closely resemble the classical definition of wet OAB. Unlike cluster F1, these women have urgency incontinence as well as urgency and frequency suggesting that OAB wet and dry are different clinical entities. This finding has been demonstrated previously in population based urodynamics studies where women with urgency incontinence were found to have maximum urethral closure pressure more similar to women with stress incontinence than OAB dry[1,2]. Participants in cluster F3 have several kinds of incontinence along with urgency and frequency. This suggests that these women might have poor urethral function given that their storage symptoms are only modest and voiding symptoms non-existent, therefore we would hypothesize that these women have poor outlet resistance consistent with SUI. Participants in cluster F4 have all LUTS including voiding, storage and incontinence reported at a severe degree suggesting that these women might have poor bladder function as well as poor outlets. Importantly, participants in this cluster present the more complex combination of symptoms than MUI, since they have equally severe levels of voiding and storage symptoms.

Previous studies have used cluster analyses to characterize women with LUTS[3,4]. Coyne et al. identified six clusters in an analysis of 8505 community dwelling women from the EPIC study on 14 lower urinary tract symptoms, including seven AUA questions. Given that study was population-based, 57% of females reported only minimal urinary symptoms. One small cluster was characterized by multiple symptoms (5%), including urinary incontinence (95%), urinary urgency (85%), terminal dribble (43%), incomplete emptying (31%), and weak stream (18%), similar to our cluster F4, which also includes patients with multiple symptoms at higher level of severity. The other clusters identified were characterized by a single symptom with the low level of other symptoms, whereas our clusters are defined by the combinations of several symptoms. This difference as well as the absence of minimal symptoms cluster in our study is likely due to differences in populations studied between LURN and EPIC. Patients in specialized urology and urogynecology clinics (LURN cohort) are likely to have a higher level of severity, but also more complicated combinations of symptoms than people with LUTS in the general population. In addition, the inclusion of the LUTS Tool in the LURN study provided higher granularity and allowed for the inclusion of symptoms that might have been missed in a shorter questionnaire.

Cluster analyses of 3167 females in the Boston Area Community Health (BACH) Survey[4] used 14 questions similar to those in EPIC study described above. Among participants, 24.1% were asymptomatic and the remainder were assigned to four clusters. The BACH clusters were largely characterized by 2–3 symptoms, with one multi-symptom cluster.

Although the studies by Coyne et al. [3] and Hall et al. [4] both clustered women from the general population and used similar questionnaires, their resultant clusters differed substantially (see detailed comparison in Rosen et al.[28]). According to Coyne et al., four of five symptomatic clusters were defined by single predominant symptoms. In contrast, Hall et al. defined all four symptomatic clusters by the combinations of symptoms; the latter result being similar to our findings. Both population studies (EPIC and BACH) and our tertiary care sample (LURN) identified a cluster in which women experienced multiple LUTS at high severity level. This cluster contained 15.2% of our cohort, 5.5% of symptomatic women in EPIC, and 8.3% of symptomatic women in BACH, which is reasonable given the fact that LURN recruited from tertiary care clinics. This cluster was found to be higher in obesity indices both in BACH and in LURN.

The strength of our subtyping methodology of LUTS patients is that we ignored the pre-conceived clinical notions and dogma surrounding LUTS in favor of utilizing purely patient-reported symptoms to derive objective clusters. Another strength is in the large sample size of the treatment-seeking cohort with all patients having physical examination, thorough medical history and demographics, as well as a voiding diary and assessment of multiple non-urologic factors in patient reported surveys. These data are available to further refine and test the subtypes of LUTS.

Our approach does carry some limitations. Clustering without clinical reasoning could result in associated but unrelated symptoms. We do not believe this to be the case given that these clusters have revealed several groups of patients that clearly distinguish themselves from one another and are common occurrences in clinical practice. The fact that conventional incontinence groups used for comparison with new clusters were based on the incontinence questions from the LUTS Tool can be conceived as a limitation of the study as well. However, we preferred this definition of the conventional groups to the one based on primary clinical impressions, which resulted in the non-mutually exclusive overlapping groups.

Our analysis only contains female subjects, and not all patients were treatment naïve. All patients were, however, naïve to treatment by the LURN physician. In addition, clustering of male patients is currently underway and the resulting subtypes will be compared with those found in the female cohort. Finally, we consider this work preliminary; the clinical significance of these subtypes is presently unknown and needs further validation. Nevertheless, we hope this clustering approach will lay the foundation for better understanding of LUTS, objective phenotyping, and personalized treatment of patients in the future.

Conclusions

Four distinct subtypes of women seeking care for LUTS identified in this study are different from community-based studies and conventional diagnostic groups. Work is currently underway to determine whether treatment and associated symptom changes over time varied by treatment group, potentially informing future cluster-specific treatments. We will continue to refine these subtypes by adding multidimensional data from other LURN studies, as well as the longitudinal symptom data (at 3 and 12 months). Future validation with the independent cohort will determine the generalizability of these clusters.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding/Support

This is publication number 8 of the Symptoms of Lower Urinary Tract Dysfunction Research Network (LURN).

This study is supported by the National Institute of Diabetes & Digestive & Kidney Diseases through cooperative agreements (grants DK097780, DK097772, DK097779, DK099932, DK100011, DK100017, DK097776, DK099879).

Research reported in this publication was supported at Northwestern University, in part, by the National Institutes of Health’s National Center for Advancing Translational Sciences, Grant Number UL1TR001422. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The following individuals were instrumental in the planning and conduct of this study at each of the participating institutions:

Duke University, Durham, North Carolina (DK097780): PI: Cindy Amundsen, MD, Kevin Weinfurt, PhD; Co-Is: Kathryn Flynn, PhD, Matthew O. Fraser, PhD, Todd Harshbarger, PhD, Eric Jelovsek, MD, Aaron Lentz, MD, Drew Peterson, MD, Nazema Siddiqui, MD, Alison Weidner, MD; Study Coordinators: Carrie Dombeck, MA, Robin Gilliam, MSW, Akira Hayes, Shantae McLean, MPH

University of Iowa, Iowa City, IA (DK097772): PI: Karl Kreder, MD, MBA, Catherine S Bradley, MD, MSCE, Co-Is: Bradley A. Erickson, MD, MS, Susan K. Lutgendorf, PhD, Vince Magnotta, PhD, Michael A. O'Donnell, MD, Vivian Sung, MD; Study Coordinator: Ahmad Alzubaidi

Northwestern University, Chicago, IL (DK097779): PIs: David Cella, Brian Helfand, MD, PhD; Co-Is: James W Griffith, PhD, Kimberly Kenton, MD, MS, Christina Lewicky-Gaupp, MD, Todd Parrish, PhD, Jennie Yufen Chen, PhD, Margaret Mueller, MD; Study Coordinators: Sarah Buono, Maria Corona, Beatriz Menendez, Alexis Siurek, Meera Tavathia, Veronica Venezuela, Azra Muftic, Pooja Talaty, Jasmine Nero. Dr. Helfand, Ms. Talaty, and Ms. Nero are at NorthShore University HealthSystem.

University of Michigan Health System, Ann Arbor, MI (DK099932): PI: J Quentin Clemens, MD, FACS, MSCI; Co-Is: Mitch Berger, MD, PhD, John DeLancey, MD, Dee Fenner, MD, Rick Harris, MD, Steve Harte, PhD, Anne P. Cameron, MD, John Wei, MD; Study Coordinators: Morgen Barroso, Linda Drnek, Greg Mowatt, Julie Tumbarello

University of Washington, Seattle Washington (DK100011): PI: Claire Yang, MD; Co-I: John L. Gore, MD, MS; Study Coordinators: Alice Liu, MPH, Brenda Vicars, RN

Washington University in St. Louis, St. Louis Missouri (DK100017): PI: Gerald L. Andriole, MD, H. Henry Lai; Co-I: Joshua Shimony, MD, PhD; Study Coordinators: Susan Mueller, RN, BSN, Heather Wilson, LPN, Deborah Ksiazek, BS, Aleksandra Klim, RN, MHS, CCRC

National Institute of Diabetes and Digestive and Kidney Diseases, Division of Kidney, Urology, and Hematology, Bethesda, MD: Project Scientist: Ziya Kirkali MD; Project Officer: John Kusek, PhD; NIH Personnel: Tamara Bavendam, MD, Robert Star, MD, Jenna Norton

Arbor Research Collaborative for Health, Data Coordinating Center (DK097776 and DK099879): PI: Robert Merion, MD, FACS; Co-Is: Victor Andreev, PhD, DSc, Brenda Gillespie, PhD, Gang Liu, PhD, Abigail Smith, PhD; Project Manager: Melissa Fava, MPA, PMP; Clinical Study Process Manager: Peg Hill-Callahan, BS, LSW; Clinical Monitor: Timothy Buck, BS, CCRP; Research Analysts: Margaret Helmuth, MA, Jon Wiseman, MS; Project Associate: Julieanne Lock, MLitt

Acknowledgements

LURN consists of six research sites and a data coordinating center. This network is conducting a prospective observational study (ClinicalTrials.gov #NCT02485808); details regarding recruitment, inclusion, and exclusion criteria have been published elsewhere.[4] The observational cohort study was approved by the Institutional Review Board at each site, and all participants provided informed consent prior to enrollment.

Abbreviation Key

AUA-SI	American Urological Association Symptom Index
BACH	Boston Area Community Health
GUPI	Genitourinary Pain Index
LURN	Symptoms of Lower Urinary Tract Dysfunction Research Network
LUTS	Lower urinary tract symptoms
MUI	mixed urinary incontinence
NIDDK	National Institute of Diabetes and Digestive and Kidney Diseases
OAB	Overactive bladder
PRO	patient reported outcome
SUI	stress urinary incontinence
UUI	urge urinary incontinence

References

1. DeLancey JO, Trowbridge ER, Miller JM, et al. Stress urinary incontinence: relative importance of urethral support and urethral closure pressure. J Urol 179;2286–90, 2008. [PubMed: 18423707]
2. DeLancey JO, Fenner DE, Guire K, et al. Differences in continence system between community-dwelling black and white women with and without urinary incontinence in the EPI study. Am J Obstet Gynecol 202:584.e1–584.e12. 2010. [PubMed: 20510959]
3. Coyne KS, Matza LS, Kopp ZS et al. Examining lower urinary tract symptom constellations using cluster analysis. BJU Int 101:1267–73, 2008. [PubMed: 18336611]
4. Hall SA, Çinar A, Link CL, et al. Do urological symptoms cluster among women? Results from the Boston Area Community Health Survey. BJU International, 101:1257–1266, 2008. [PubMed: 18419699]
5. Michel MC, Oelke M, Goepel M, et al. Relationships among symptoms, bother, and treatment satisfaction in overactive bladder patients. Neurourol Urodyn 26:190–5, 2007. [PubMed: 17096320]
6. Yang CC, Weinfurt KP, Merion RM, et al. Symptoms of Lower Urinary Tract Dysfunction Research Network. J Urol 196:146–52, 2016. [PubMed: 26784646]
7. Cameron AP, Lewicky-Gaupp C, Smith AR et al. Baseline Lower Urinary Tract Symptoms in Patients Enrolled in LURN: A Prospective, Observational Cohort Study. J Urol 199:1023–1031, 2018. [PubMed: 29111381]
8. Coyne KS, Sexton CC, Kopp Z, et al. Assessing patients' descriptions of lower urinary tract symptoms (LUTS) and perspectives on treatment outcomes: results of qualitative research. Int J Clin Pract 64:1260–78, 2010. [PubMed: 20579138]
9. Coyne KS, Barsdorf AI, Thompson C, et al. Moving towards a comprehensive assessment of lower urinary tract symptoms (LUTS). Neurourol. Urodyn 31:448–454, 2012. [PubMed: 22396308]
10. Barry M, Fowler F, Jr, O'Leary M et al.: The American Urological Association symptom index for benign prostatic hyperplasia. The Measurement Committee of the American Urological Association. J Urol 1992; 148: 1549. [PubMed: 1279218]
11. Barry M, Fowler F, Jr, O'Leary M et al.: Measurement Committee of the American Urological Association. Med Care 1995; 22: AS145.
12. Spiegel BM, Hays RD, Bolus R, et al. Development of the NIH Patient-Reported Outcomes Measurement Information System (PROMIS) Gastrointestinal Symptom Scales. Am. J. Gastroenterol 109:1804–14, 2014 PMID: 25199473 [PubMed: 25199473]
13. Pilkonis PA, Choi SW, Reise SP, et al. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS®): depression, anxiety, and anger. Assessment 18:263–83, 2011. [PubMed: 21697139]
14. Cohen S, Kamarck T, Mermelstein R: A global measure of perceived stress. J. Health Soc. Behav 24:385–96, 1983. [PubMed: 6668417]
15. Yu L, Buysse DJ, Germain A, et al.: Development of short forms from the PROMIS™ sleep disturbance and Sleep-Related Impairment item banks. Behav Sleep Med 10: 6–24, 2011. [PubMed: 22250775]
16. Clemens JQ, Calhoun EA, Litwin MS, et al. Validation of a Modified National Institutes of Health Chronic Prostatitis Symptom Index to Assess Genitourinary Pain in Both Men and Women. Urology 74:983–7, 2009. [PubMed: 19800663]
17. Barber MD, Chen Z, Lukacz E, et al. Further validation of the short form versions of the Pelvic Floor Distress Inventory (PFDI) and Pelvic Floor Impact Questionnaire (PFIQ). Neurourol. Urodyn 30:541–546, 2011. [PubMed: 21344495]
18. Raghunathan TE, Lepkowski JM, Van Hoewyk J, et al. A multivariate technique for multiply imputing missing values using a sequence of regression models. Survey Methodol 27:85–95, 2001.
19. Raghunathan TE, Solenberger PW, Berglund P, et al. IVEware: Imputation and variance estimation software Ann Arbor: University of Michigan, Institute for Social Research, Survey Research Center 2000.
20. Monti S, Tamayo P, Mesirov J, et al. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. Mach Learn 52:91–118, 2003.

21. Davies DL, Bouldin DW. A Cluster Separation Measure. IEEE Trans Pattern Anal Mach Intell 1:224–227, 1979. [PubMed: 21868852]
22. Dunn JC. A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. J Cybernetics 3:32–57, 1973.
23. Pal NR, Biswas J. Cluster validation using graph theoretic concepts. Pattern Recognit 30:847–857, 1997.
24. Baker FB, Hubert LJ. Measuring the Power of Hierarchical Cluster Analysis. J Am Stat Assoc 70:31–38, 1975.
25. Milligan GW. A monte carlo study of thirty internal criterion measures for cluster analysis. Psychometrika 46:187–99, 1981.
26. Calinski T, Harabasz J. A dendrite method for cluster analysis. Commun Stat 3:1–27, 1974.
27. Liao TW. Clustering of time series data -a survey. Pattern Recognit 38:1857–1874, 2005.
28. Rosen RC, Coyne KS, Henry D, et al. Beyond the cluster: methodological and clinical implications in the Boston Area Community Health survey and EPIC studies. BJU Int 101:1274–8, 2008. [PubMed: 18419700]

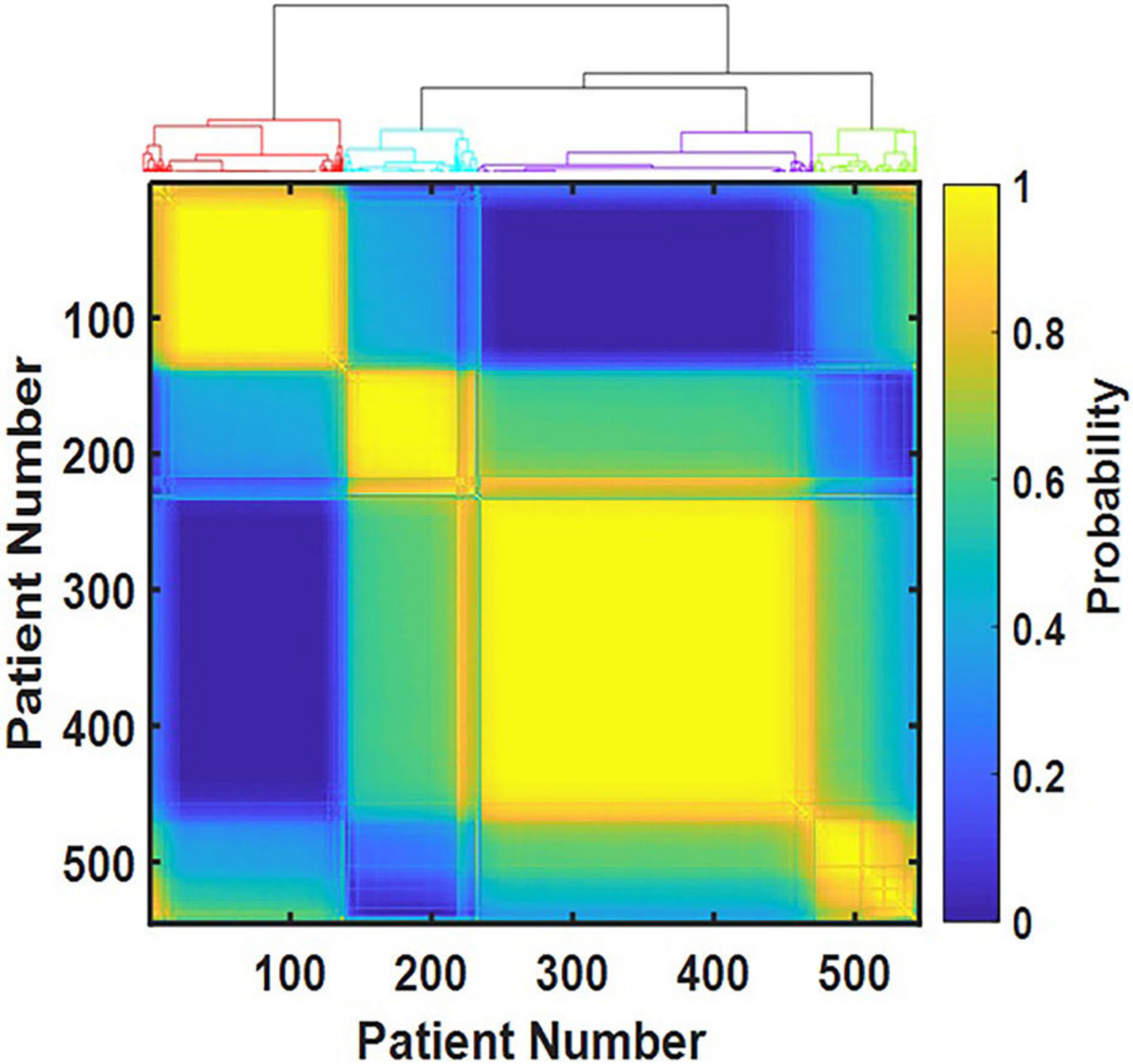

Figure 1.
Consensus matrix as a color map. Each element of the 545 by 545 matrix represents the probability that the respective pair of participants both belong to the same cluster. Probability is color-coded: bright yellow represents probability close to one, dark blue – probability close to zero. Four bright yellow squares along the diagonal represent four identified clusters of participants. The dendrogram on top of the consensus matrix demonstrates four distinct clusters as well.

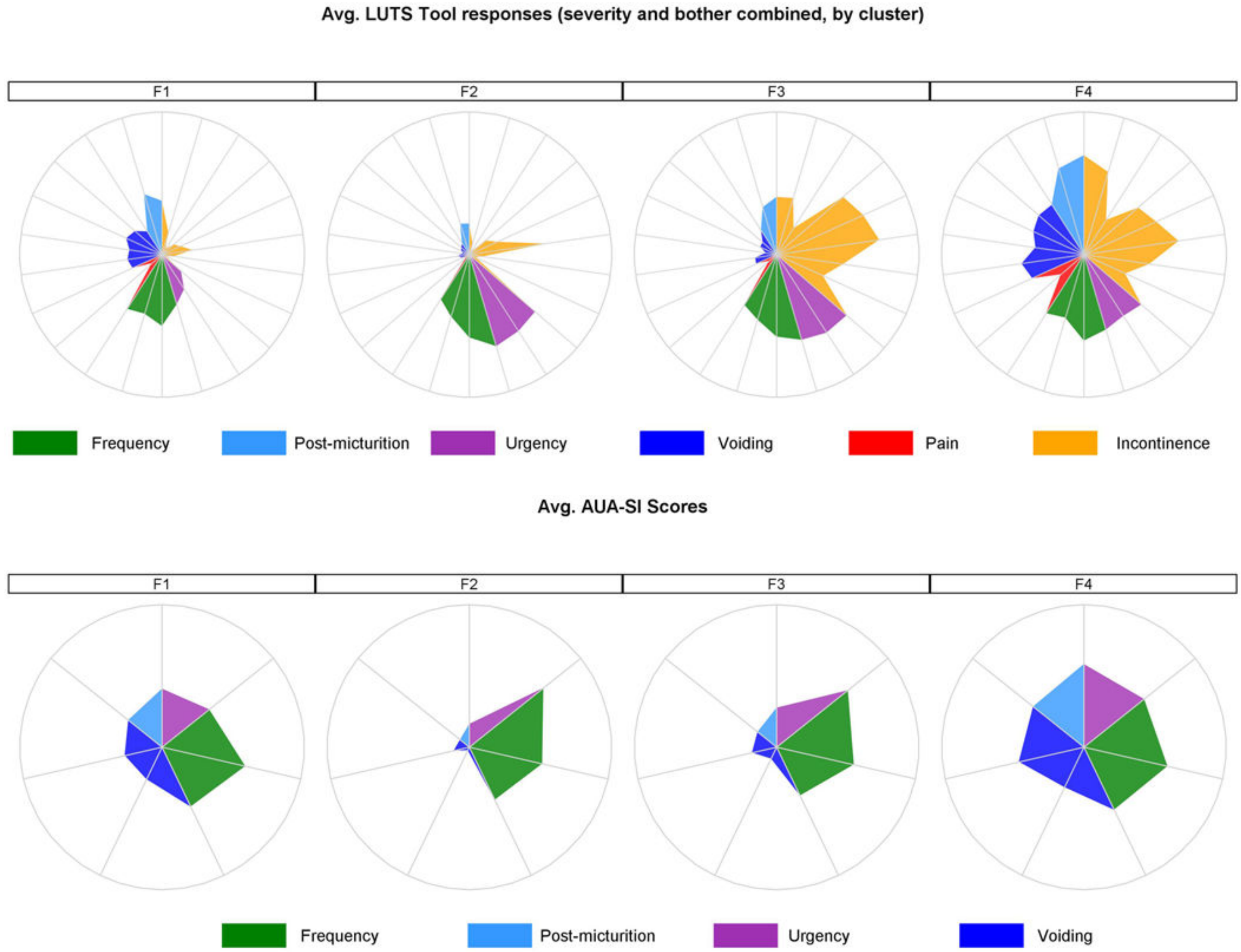

Figure 2.

Symptom signatures of four clusters as radar plots. Figure 2a represents signatures based on the LUTS tool questions, where each "spoke" is an average of symptoms' severity and bother for the females within the given cluster. Figure 2b represents averaged answers within the cluster to the AUA-SI questions. Questions related to similar symptoms are grouped together and color-coded. LUTS questions are about: 1-frequency, 2-daytime frequency, 3-nocturia, 4-incomplete emptying, 5-trickle/dribble, 6-urgency, 7-hesitancy, 8-intermittency, 9-strain, 10-weak stream, 11-splitting/spraying, 12-urgency w/fear, 13-pain, 14-burning, 15-leakage, 16a-leakage post voiding, 16b-leakage w/urgency, 16c-leakage w/laugh, 16d-leakage w/exercise, 16f-leakage w/sex, 16g-leakage for no reason. AUA questions are about: 1-nocturia, 2-incomplete emptying, 3-frequency, 4-intermittency, 5-urgency, 6-weak stream, 7-strain. Circles represent the highest severity level (typically answer #5 to each question).

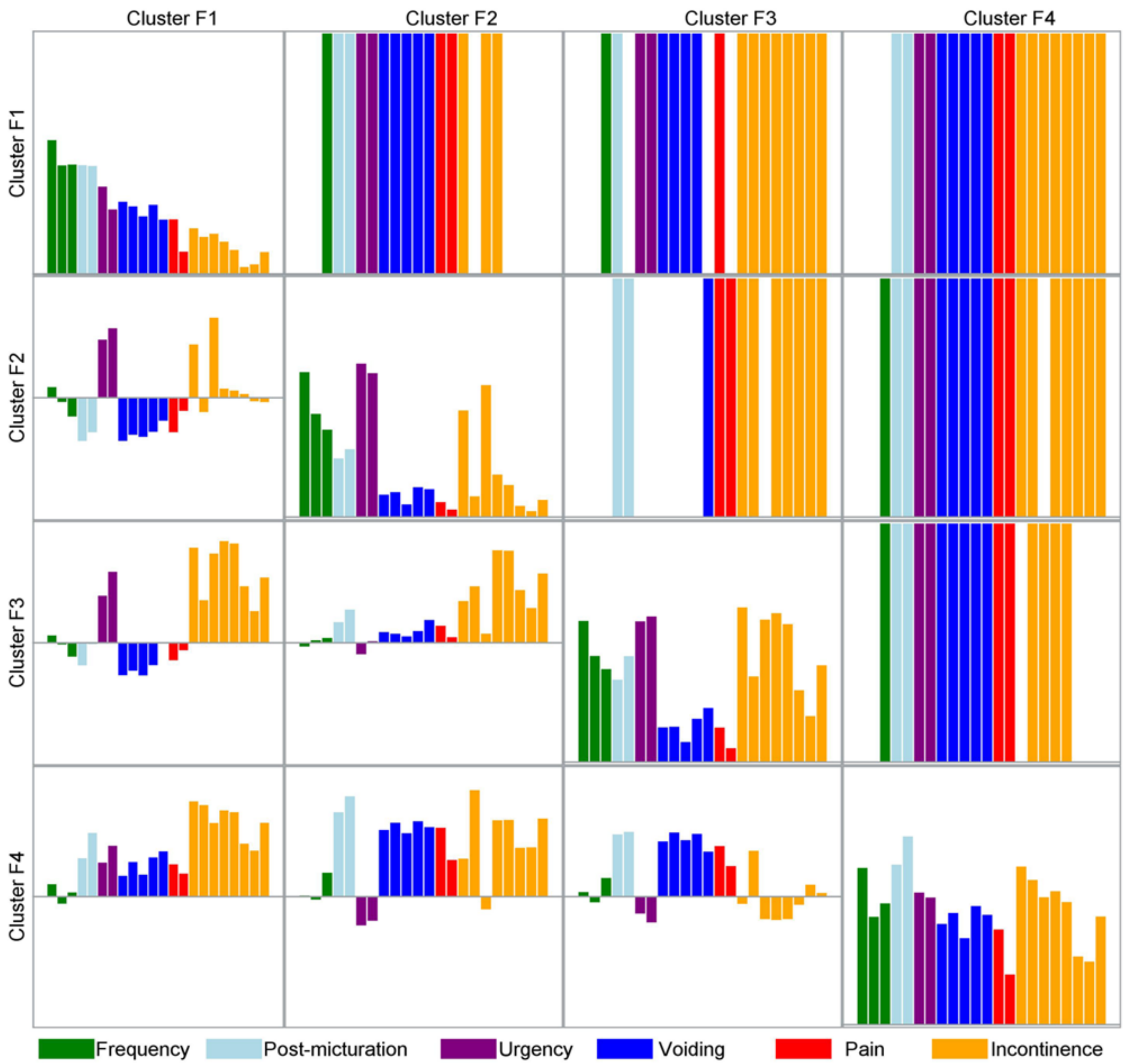

Figure 3.

Mean symptom severity levels by cluster, pairwise symptom severity differences between clusters and statistical significance. On-diagonal rectangles represent mean severity level for each of the LUTS tool symptoms for each cluster. Colored bars in the upper triangle of the matrix represent symptoms that are significantly different in each possible pair of four clusters. For instance, the second rectangle in the first row demonstrates that 14 symptoms are significantly different in cluster F1 versus cluster F2: one related to frequency, two to post-micturition, two to urgency, five to voiding, two to pain, and two to incontinence. The elements in the lower triangle of the matrix present the difference in the symptom severity levels between clusters; e.g., the first (upper) element in the triangle represents the difference between symptom severity levels in cluster F2 and cluster F1, indicating that urgency symptoms are more severe in cluster F2, while voiding, storage, and pain symptoms are more severe in cluster F1. Similarly, matrix in Supplemental Figure 2 is based on the bother level for each of LUTS Tool questions, while matrix in Supplemental Figure 3 is based on the AUA-SI questions.

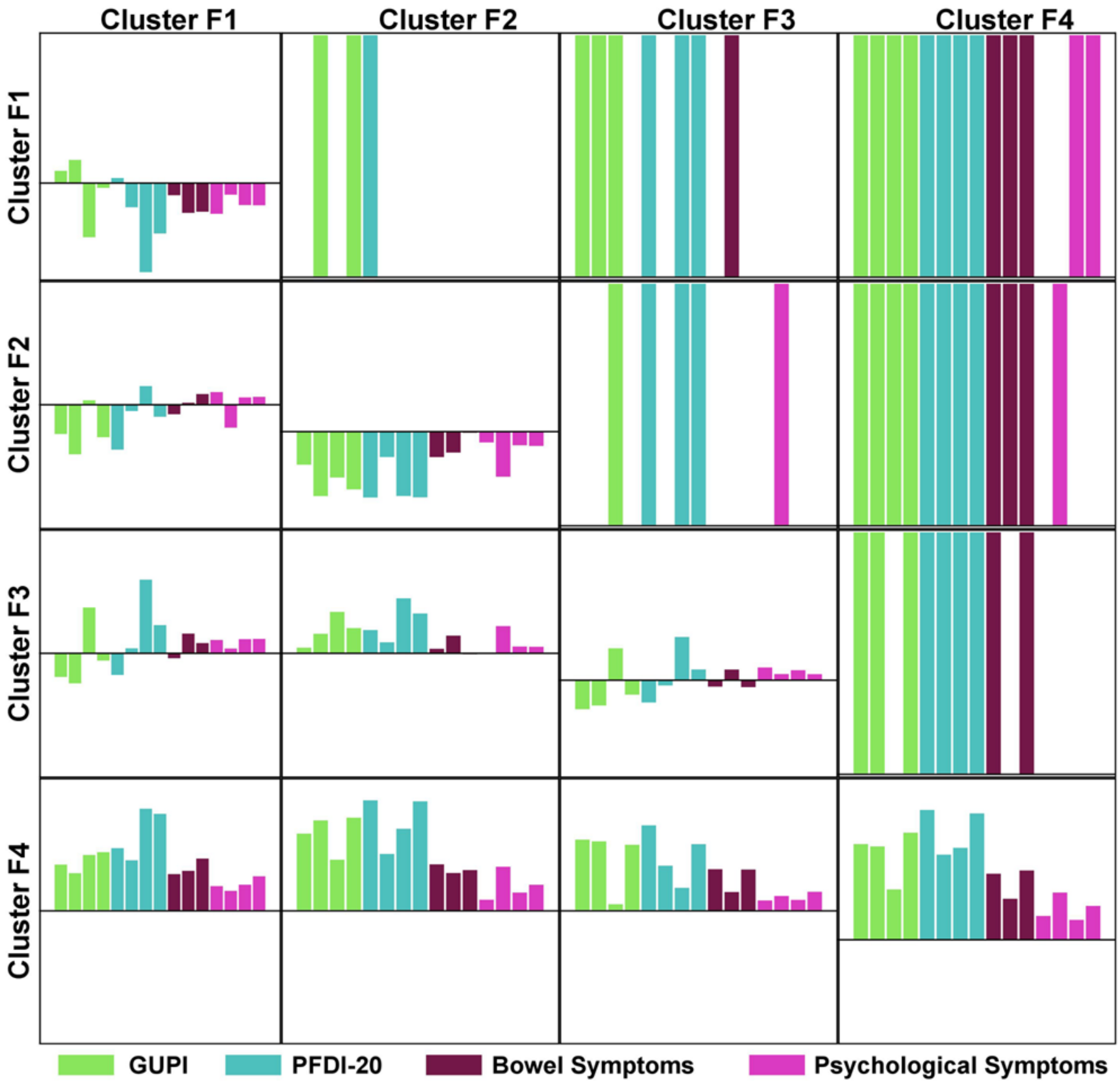

Figure 4.
Non-urologic patient reported outcomes (PRO). Matrix of pairwise comparisons of the clusters based on adjusted regression models. Similarly to Figure 3, upper triangle of the matrix presents PROs significantly different in adjusted pairwise comparison of the clusters. Lower triangle presents adjusted estimated differences in PROs for each pair of clusters. On-diagonal rectangles represent mean scores for each PRO for each cluster. Note that since the PROs had varying scales, all scores were converted to Z-scores with mean 0 and variance 1, therefore negative values of PROs indicate that the average value for the cluster was below the overall mean score for a given PRO. PROs are: 1-GUPI Pain Subscale, 2-GUPI Urine Subscale, 3-GUPI QOL Subscale, 4-GUPI Total Score, 5-PODI-6, 6-CRADI-8, 7-UDI-6, 8-PFDI-20,9-PROMIS Constipation, 10-PROMIS Diarrhea, 11-PROMIS Bowel Incontinence,

12-Perceived Stress Scale, 13-PROMIS Sleep Dysfunction, 14-PROMIS Depression, 15-PROMIS Anxiety.

Author Manuscript
Author Manuscript
Author Manuscript

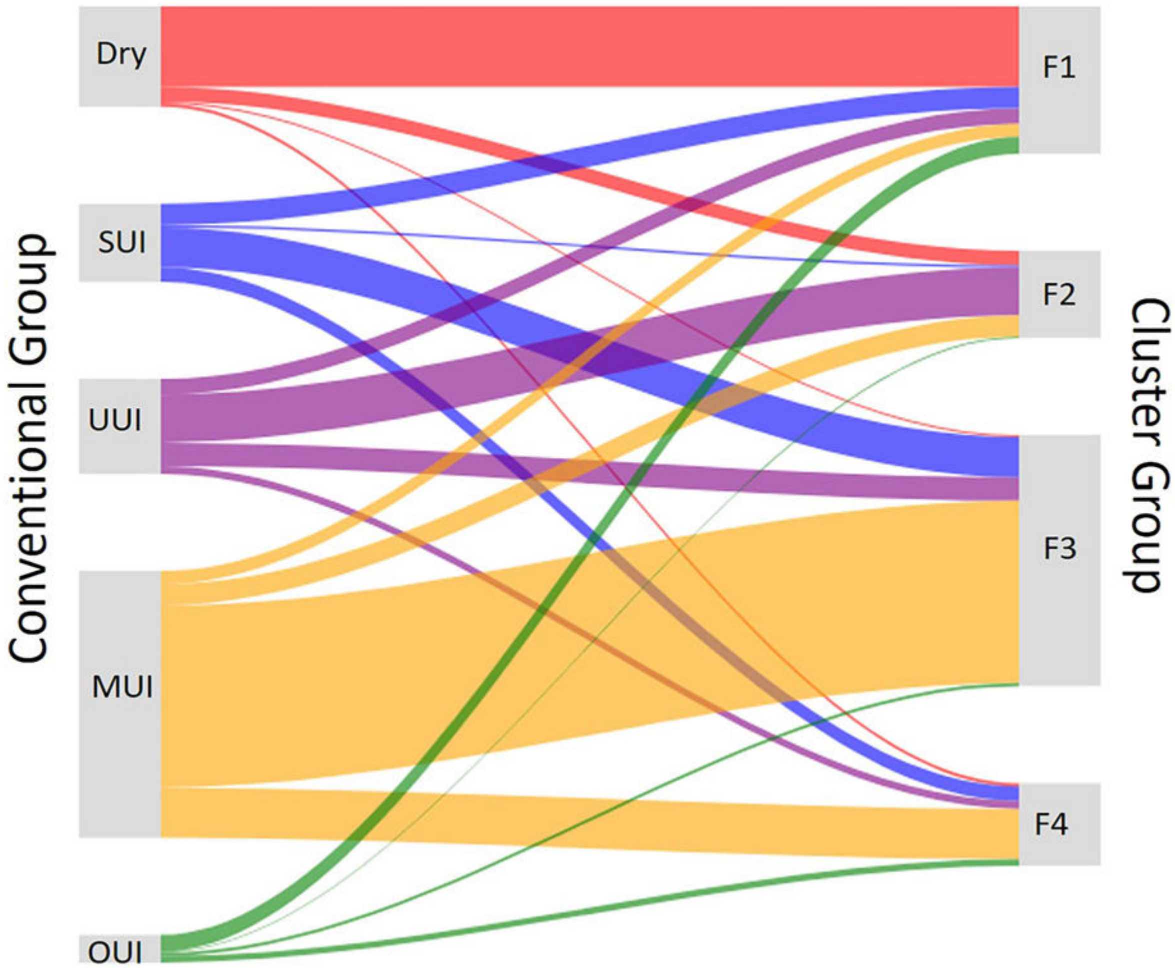

Figure 5.

Sankey diagram comparison of group membership. Conventional groups (on the left): dry (continent), stress urinary incontinent, urgency urinary incontinent, mixed urinary incontinent, and other urinary incontinent are compared with four clusters identified in this paper (on the right). Each of conventional groups contributes members to each of the new clusters. “Dry” means continent, SUI-stress urinary incontinent, UUI- urge urinary incontinent, MUI- mixed urinary incontinent, OUI- other urinary incontinent.

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Demographic Data for Each of the Clusters

	Cluster F1	Cluster F2	Cluster F3	Cluster F4	P-Value
N	138	80	244	83	
Age	55.8	58.5	56.5	55.3	0.4878
Race					0.5507
White	113(82%)	65(81%)	202(83%)	65(78%)	
Black	13(9%)	10(13%)	30(12%)	11(13%)	
Asian	4(3%)	2(3%)	6(3%)	2(2%)	
American Indian	3(2%)	0(0%)	0(0%)	2(2%)	
Native Hawaiian	0(0%)	1(1%)	0(0%)	0(0%)	
Other	5(4%)	2(3%)	5(2%)	3(4%)	
Ethnicity					0.1967
Non-Hispanic/latino	131(95%)	79(99%)	224(92%)	78(94%)	
Hispanic or Latino	4(3%)	1(1%)	14(6%)	2(2%)	
Unknown	3(2%)	0(0%)	6(3%)	3(4%)	
Obese	42(30%)	32(40%)	134(55%)	39(47%)	<0.0001
Post-Menopausal	83(60%)	57(71%)	150(62%)	57(69%)	0.5897
Had a Hysterectomy	38(28%)	22(28%)	70(29%)	34(41%)	0.1271
At least one Vaginal Birth	87(63%)	54(68%)	185(76%)	63(76%)	0.0308
Alcoholic Drinks per Week					0.3835
Never	23(17%)	8(10%)	40(16%)	19(23%)	
0–3 Drinks Per Week	92(67%)	57(71%)	158(65%)	48(58%)	
4–7 Drinks Per Week	15(11%)	9(11%)	31(13%)	11(13%)	
More than 7 Drinks Per Week	2(1%)	2(3%)	11(5%)	3(4%)	
Smoking Status					0.2888
Never Smoker	94(68%)	58(73%)	149(61%)	47(57%)	
Former Smoker	34(25%)	19(24%)	72(30%)	28(34%)	
Current Smoker	6(4%)	3(4%)	20(8%)	6(7%)	
Smoking Status Unknown	4(3%)	0(0%)	3(1%)	2(2%)	