Clinically Important Differences for Pain and Urinary Symptoms in Urologic Chronic Pelvic Pain Syndrome: A MAPP Network Study

Alisa J. Stephens-Shields[1,*], **H. Henry Lai**[2], **J. Richard Landis**[1], **Karl Kreder**[3], **Larissa V. Rodriguez**[4], **Bruce D. Naliboff**[5], **Niloofar Afari**[6], **Siobhan Sutcliffe**[2], **Robert Moldwin**[7], **James W. Griffith**[8], **J. Quentin Clemens**[9], **Catherine S. Bradley**[3], **Susan Quallich**[9], **Priyanka Gupta**[9], **Steven E. Harte**[9], **John T. Farrar**[1], **MAPP Research Network**

[1]University of Pennsylvania, Philadelphia, PA

[2]Washington University, St. Louis, MO

[3]University of Iowa, Iowa City, IA

[4]Weill Cornell Medicine, New York, NY

[5]University of California, Los Angeles, CA

[6]University of California, San Diego, CA

[7]Zucker School of Medicine at Hofstra-Northwell, Lake Success, NY

[8]Northwestern University, Chicago, IL

[9]University of Michigan, Ann Arbor, MI

Abstract

Objectives: Symptom heterogeneity in interstitial cystitis/bladder pain syndrome and chronic prostatitis/chronic pelvic pain syndrome, collectively termed urologic chronic pelvic pain syndrome (UCPPS), has resulted in difficulty in defining appropriate clinical trial endpoints. We determine clinically important differences (CIDs) for two primary symptom measures, Pelvic Pain Severity (PPS) and Urinary Symptom Severity (USS), and evaluate subgroup differences.

Methods: The Multidisciplinary Approach to the Study of Chronic Pelvic Pain Symptom Patterns Study enrolled individuals with UCPPS. We defined CIDs by associating changes in PPS and USS over 3 to 6 months with marked improvement on a global response assessment using regression and receiver operating characteristic curves. We evaluated CIDs for absolute and percent change and examined differences in CIDs by sex-diagnosis, presence of Hunner lesions, pain type, pain widespreadness, and baseline symptom severity.

Results: An absolute change of −4 was clinically important in PPS among all patients, but CID estimates differed by pain type, presence of Hunner lesions, and baseline severity. PPS CID estimates for percent change were more consistent across subgroups and ranged from 30 to 57 percent. The absolute change USS CID was −3 for female participants and −2 for male

*Corresponding Author: alisaste@pennmedicine.upenn.edu.

participants with CP/CPPS only. Patients with greater baseline severity required larger decreases in symptoms to feel improved. Estimated CIDs had lower accuracy among participants with low baseline symptoms.

Conclusion: A reduction of 30–50% in PSS is a clinically meaningful endpoint for future therapeutic trials in UCPPS. USS CIDs are more appropriately defined separately for male and female participants.

INTRODUCTION

Pelvic pain and urinary symptoms are cardinal features of interstitial cystitis/bladder pain syndrome (IC/BPS) and chronic prostatitis/chronic pelvic pain syndrome (CP/CPPS). Collectively termed urologic chronic pelvic pain syndromes (UCPPS), both populations are phenotypically heterogeneous. This heterogeneity has engendered tremendous difficulties when selecting outcomes for therapeutic clinical trials. Ideal outcomes are validated, reliable, and interpretable. Identification of clinically important differences (CIDs) [1,2], changes in symptoms that patients perceive as meaningful, increases interpretability of trial outcomes. A novel treatment may be deemed effective if it leads to more patients improving symptoms by the amount of the CID than a standard treatment or placebo.

Using longitudinal data from The Multidisciplinary Approach to the Study of Chronic Pelvic Pain (MAPP) Research Network, we determined CIDs for the Pelvic Pain Severity (PPS) index and the Urinary Symptom Severity (USS) index[3] for future use as endpoints in UCPPS therapeutic trials. These outcomes, derived from the Genitourinary Pain Index (GUPI) pain and urinary subscores and the Interstitial Cystitis Symptom Index, demonstrated reliability and discriminant and convergent validity in men and women with UCPPS[4]. Even though PPS and USS are moderately correlated(r=0.53), they show unique patterns of relationships with other measures and different longitudinal trajectories[5], which supports separate measurement and evaluation in future studies[3]. We additionally determined CIDs for the Brief Pain Inventory Item 6 (BPI-6), and the GUPI Item 4 (GUPI-4), two 0–10 pain scales, to compare with 0–10 outcomes used for other pain conditions. Finally, we assess differences in CIDs by UCPPS subgroups.

METHODS

Sample

The Multidisciplinary Approach to the Study of Chronic Pelvic Pain (MAPP) Research Network Symptom Patterns Study (SPS: MAPP II)[6] recruited 620 patients with symptoms of UCPPS from July 2015 to February 2019. Eligible individuals were 18 years old; had feelings of pain, pressure, or discomfort in the pelvic area for the majority of the last 3 months; reported pain, pressure or discomfort of at least 1 on the 0–10 scale for UCPPS symptoms during the past 2 weeks; and had a diagnosis of IC/BPS or CP/CPPS or confirmation of UCPPS symptoms by a knowledgeable clinician. Returning patients from MAPP I were allowed to enroll with a score of 0 to provide insight into symptom resolution. Complete eligibility and exclusion criteria have been described[6]. A total of 578 participants completed a 4-week screening and were followed for 3 years, with semi-annual in-clinic

assessments and quarterly internet-based assessments of urologic symptoms, neuroimaging, quantitative sensory testing, and biomarker collection. Our analytic sample included MAPP participants with PSS or USS at the initial visit and either month 3 or 6 to enable defining change from baseline, and a GRA in the same month as change from baseline(n=537). Forty-one participants were excluded for missing change in the PPS and USS or the GRA at both 3 and 6 months. There were 536 and 535 participants contributing to the analysis for PPS and USS, respectively.

Measures

The PPS index (0–28) indicates an individual's severity of pain in the pelvis and bladder, with slightly different questions based on male and female anatomy[3,4]. The USS index (0–25) indicates an individual's severity of urinary symptoms, including frequency and urinary urgency[3,4]. Items and scoring are available online[4]. Reliability of PPS and USS were 0.76 and 0.86, respectively[4]. Assessments of construct validity demonstrated adequate association with pelvic pain and urinary symptom measures. Complete psychometric analysis has been described[3,4]. We further estimated CIDs of two 0–10 items: item 4 of the GUPI (GUPI-4) and item 6 of a tailored Brief Pain Inventory (BPI-6) for UCPPS patients[6]. GUPI-4 assesses participants' average pelvic pain or discomfort on the days that it was experienced over the last week. BPI-6 asks participants to rate their average pain (not limited to pelvic region) over the last week.

The Global Responses Assessment (GRA) was used as the gold-standard responder status to anchor symptom measures in estimating CIDs. The GRA is a numerical response scale that asks participants, 'As compared to when you started the study, how would you rate your overall symptoms now?' with 'markedly worse', 'moderately worse', 'slightly worse', 'no change', 'slightly improved', 'moderately improved', or 'markedly improved' as responses.

We examined CID heterogeneity by sex-diagnosis (male with CP/CPPS only, male with IC/BPS [also could have CP/CPPS], or female), presence or absence of Hunner lesions, likely having a neuropathic component or non-neuropathic phenotype using the painDetect scale[7] (painDetect > 12 likely neuropathic; painDetect   12 likely non-neuropathic), widespread or localized pain (fibromyalgianess scale[8,9] > 7 or <=7), whether the most bothersome symptom was pain- or urinary-related (PPS Bother or USS Bother), and baseline tertile of the respective symptom. Detection of Hunner lesion (presence or absence) by cystoscopy was described previously[10].

Statistical Methods

To assess the quality of the GRA as the anchor, we first assessed correlation of the GRA with concurrent absolute change at month 3 or 6 (symptom score at month 3 or 6 minus symptom score at baseline), percent change (absolute change divided by baseline score x 100%), and month 3 or 6 score of PPS, USS, BPI-6, and GUPI-4 using Spearman's rank correlation. We also determined correlations among the four outcomes. We then used linear regression models and responder operating characteristic (ROC) curves to estimate CIDs for PPS, USS, BPI-6 and GUPI-4. Regression models included absolute change as the outcome and categorical 7-level GRA as the primary predictor; models adjusted for month (3 or 6).

Models were fit using generalized estimating equations (GEE) with working independence correlation[11] (Appendix 1). We marginalized model coefficients over month to obtain mean change by GRA category and adopted mean change for the 'markedly improved' group as the regression-based estimated CID. We also report mean change for the 'moderately improved'. To evaluate differences by subgroups, we added subgroup-defining variables and their two-way interaction with GRA to models and tested the equality of predicted marginal means of change for marked improvement in each subgroup. We also fit models with percent change as the outcome, and models adjusted for the baseline score of the respective symptom to determine if baseline scores explained differences in CID estimates across subgroups.

We also estimated CIDs by calculating sensitivity, specificity, and corresponding receiver operating characteristic (ROC) curves. We defined a clinically important responder as 'markedly improved' versus other on the GRA and calculated sensitivity and specificity based on counts of individuals in the 2×2 cross classification of dichotomized GRA and achieving or not achieving a candidate threshold of change in the symptom score. We then generated ROC curves and area under the curve (AUC). At each candidate CID, we calculated the minimum between sensitivity and specificity and selected as the CID the value maximizing that minimum across candidate CIDs.[12] We examined heterogeneity by repeating analyses within subgroups. Secondary analyses defined a clinically important responder as 'markedly' or 'moderately improved'. Parallel analyses considered the clinically meaningful score achieved at follow-up and percent change. As a secondary analysis, we repeated regression and ROC analyses among participants with baseline GUPI-4 ≥ 4 to reflect an active state of disease. Analyses were performed in Stata 16 and R 3.6.3.

RESULTS

A total of 536 MAPP II SPS participants contributed 960 observations across months 3 and 6; 42(7.1%) participants did not report PPS symptoms or accompanying GRA responses at 3 or 6 months and did not contribute to the analysis. Similarly, 535 participants had available USS data. The median (IQR) age was 44(31,58) years, and 385 (67%) were female (Table 1). Of 193 males, 126 (65%) had CP/CPPS only, 23 (12%) had IC/BPS only, 35 (18%) had both conditions, and 9(5.0%) had neither prior diagnosis but reported symptoms consistent with IC/BPS. The majority of participants had a likely non-neuropathic phenotype (363(64%)), and 297(51%) had widespread non-pelvic pain. At baseline, pelvic pain (versus a urinary symptom) was reported as most bothersome for 397 (72%) participants. The median PPS, USS, BPI-6, and GUPI-4 at baseline were 16(12,19), 13(7,17), 4 (3,6), and 4 (2,5), respectively. The correlations of the GRA with change in symptoms at month 3 were 0.43, 0.42, and 0.40 for PPS, BPI-6, and GUPI-4, respectively and 0.28, for USS (Table 2). Correlations of the GRA with percent change and score were similar at months 3 and 6 (data not shown). At 3 and 6 months, 34 (7.1%) and 41 (8.7%) of participants reported 'marked improvement', respectively (Table 1). There were 383 of 536 (71%) participants with GUPI-4 ≥ 4 at baseline.

Pelvic Pain Severity (PPS)

The regression-based mean (95% CI) change from baseline in PPS, marginalized across visits, was −6.8 (−7.7, −5.9) for 'markedly improved' participants and −4.3 (−5.0, −3.7) for 'moderately improved' (Figure 1a, Appendix Table 1). The mean PPS change by GRA response category did not differ across sex-diagnosis, pain widespreadness, or PPS/USS Bother. 'Markedly improved' participants with a likely neuropathic phenotype had a PPS mean change of −10.1 (−12.1, −8.1), whereas those with a likely non-neuropathic phenotype had a mean change in PPS of −5.6 (−4.6, −6.7) (p<0.001). Participants with Hunner lesions required a greater reduction in PPS to feel better (Figure 2, Appendix Table 2). Change in PPS from baseline varied by baseline PPS tertile, with CIDs increasing in magnitude with greater baseline severity (p<0.001). Differences in estimated CIDs by pain type and presence of Hunner lesions were reduced and not statistically significantly for percent change (p=0.1 pain type, p=0.4 Hunner lesions). Differences in CID estimates by tertile remained with percent change (p=0.01). Among patients with GUPI-4   4 at baseline the mean PPS change among 'markedly improved' participants was −8.7(−9.8, −7.5). The CID estimate based on percent change for the subset with baseline GUPI-4   4 (−48.9% (95% CI −55.7%, −42.1%) was similar to the percent change CID estimate in the full sample −43.1% (−52.2%, −34.1%).

ROC analysis resulted in an estimated CID of −4 for PPS (AUC=0.75) overall(Table 3). Across subgroups, ROC-based CID estimates ranged from −9 to −1. The clinically important score of PPS achieved was 10 (AUC=0.8), ranging from 7 to 14 by subgroups. AUCs were higher using the more specific 'markedly improved' GRA outcome versus at least 'moderately improved' and for analyses of PPS score achieved. Considering percent change, the estimated CID was −30% overall. Estimated CIDs were generally 1 point greater in magnitude when restricting the sample to patients with baseline GUPI-4   4 (Appendix Table 4).

Urinary Symptom Severity (USS)

Markedly and moderately improved patients had mean (95% CI) USS changes of −3.5(−4.2, −2.8) and -2.1(−2.6,−1.5), respectively (Figure 1, Appendix Table 1). 'Markedly improved' female participants reported greater decreases in USS than 'markedly improved' male participants of either pain-only or urinary symptom phenotypes (Figure 2, Appendix Table 2, p<0.001). No significant differences were observed in estimates of mean USS change by GRA response across other subgroups, except for baseline USS tertile (Figure 2, Appendix Table 2). Accounting for differences in baseline USS by regression adjustment or using percent change did not explain differences by sex-diagnosis or USS tertile (Figure 2). Restricting to patients with GUPI-4   4 did not appreciably change USS CID estimates (Appendix Table 2).

In the overall sample, ROC analysis indicated a change of −2 for the estimated CID in USS (AUC=0.68) and a range of −4 to −1 across subgroups (Table 3). A clinically meaningful USS score achieved was 8 (AUC=0.69), ranging from 6 to 11 across subgroups. Classification accuracy between change and score achieved was similar.

Brief Pain Inventory Item 6 and Genitourinary Pain Index Item 4

Appendix Figure 1 displays regression-based CID estimates for absolute and percent change for BPI-6 and GUPI-4. Appendix Table 3 summarizes ROC-based CIDs. For BPI-6, we observed differences in regression-based CID estimates for likely neuropathic versus likely non-neuropathic phenotypes, but not by sex-diagnosis, Hunner lesions, pain widespreadness, or most bothersome symptom type. The ROC-based threshold in the overall sample was −2 (AUC=0.68) and ranged from −3 to −1 across subgroups. CID estimates (range across subgroups) were −30% (−35% to −5%) and 2 (1 to 4) for percent change and score, respectively. For GUPI-4, there were no differences in the mean change among improved participants by subgroups except for baseline tertile. The ROC-derived CID estimate was −2 overall and in most subgroups. The ROC-based CID estimate for percent change was −35% overall and −55% to −30% across subgroups. The clinically important score achieved was 2 overall, ranging from 1 to 3 across subgroups. Restriction to participants with GUPI-4 4 at baseline increased regression-based GUPI-4 CID estimates by approximately 1 point in magnitude; ROC-based CID estimates were similar to the full sample (Appendix Table 5).

DISCUSSION

We quantify meaningful improvement in UCPPS by estimating CIDs for several UCPPS symptom measures. These CIDs represent evidence-based endpoints for future clinical trials. We further evaluated differences in CIDs by subgroups considering absolute change, percent change, and score achieved using multiple analytic approaches.

We determined absolute changes of −6.8 and −4.0 in the 28-point PPS, and −3.5 and −2.0 in the 25-point USS to be clinically meaningful in the overall sample using regression and ROC methods, respectively, and 'marked improvement' as the anchored definition of a responder. We found −2.6 and −2.0 to be meaningful for GUPI-4 and −2.0 and −1.8 for BPI-6. Our CID estimates for average general pain (BPI-6) are consistent with values in other pain syndromes[13]. Estimated CIDs based on the more sensitive definition of at least moderate improvement were modestly attenuated and had lower classification accuracy.

We observed subgroup variation in CIDs. Female participants and males with CP/CPPS only required larger absolute and percent decreases in urinary symptoms to feel better than males with urinary symptoms. Larger absolute improvement was necessary for participants in both urologic and general pain symptoms to feel improved if they were likely to have a neuropathic pain phenotype. Participants with Hunner lesions required larger absolute improvement in urologic pain to feel improved than participants without Hunner lesions. Given the differences pathophysiology, treatment approaches and CIDs between patients with and without Hunner lesions, it would be appropriate to assess these two groups separately in future clinical trials. Across all outcomes, participants with greater symptom severity at the start of the study required greater absolute and percent changes to feel improved. Differences were also observed, however, across subgroups in the score achieved for clinically meaningful improvement, thereby complicating the adoption of a single clinically important endpoint based on follow-up symptom scores. Comparison of CID estimates by baseline symptom tertile often revealed similar estimates of absolute and

percent change among middle and high tertiles, but substantially lower estimates of change among the low baseline tertile, with lower accuracy likely due to floor effects.

Our findings highlight a challenge of the analysis of patients with modest symptom severity in efficacy studies. In the low tertile, estimates of clinically meaningful change were not distinguishable from random variation, casting doubt on symptom reductions as real improvement. For trials that include participants with modest symptoms, defining the endpoint based on the score achieved after therapy may be more appropriate but still subject to higher misclassification error than for participants starting with more severe symptoms. In such case, it would be prudent to account for differences in baseline values. Alternatively, UCPPS patients with minimal/moderate symptoms noted at study inception could be excluded.

Absolute change, percent change, and follow-up score achieved may be dichotomized for interpretability, which may be prioritized over the higher power of a continuous endpoint. Classification error may point to one endpoint dichotomization over another in defining clinically meaningful endpoints. In our study, the follow-up score had better accuracy than absolute or percent change, but in modestly sized studies, chance imbalances in baseline severity may impact inferences based solely on follow-up score. Regardless of the dichotomous endpoint selected, we recommend always examining continuous change or post-intervention scores to fully characterize the distributions of responses among patients allocated to alternate treatments.

This study has several strengths, including a large sample, protocol-specified time intervals for data collection, strong correlation of the anchor with PPS, and little missing data. Limitations include the lower correlation of the GRA with USS and reliance on a single anchor across urologic pain, general pain, and urinary symptoms. Additionally, although our sample was large, it included participants recruited from tertiary care centers and subspecialty clinics and may not reflect the general population of UCPPS. However, because our primary comparison was between scales reported by individual patients, our findings represent reasonable estimates of the appropriate CID for the scales tested.

CONCLUSION

Our findings provide a comprehensive approach to the study of CIDs and score achieved in two primary symptom areas of UCPPS, urinary and pelvic pain. Differences in CIDs across subgroups challenged threshold selection, but restriction to mid and high baseline severity patients and use of percent change increased consistency of CIDs across subgroups; 30%–50% reductions in PPS were identified as clinically meaningful depending on regression or ROC analysis. USS CIDs differed by sex-diagnosis. Restricting to participants with pelvic pain 4 on a 0–10 scale at baseline increased the magnitude of absolute, but not percent clinically important differences for PSS. CID estimates among low severity patients were more prone to error.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Dworkin RH, Turk DC, McDermott MP, Peirce-Sandner S, Burke LB, Cowan P, Farrar JT, Hertz S, Raja SN, Rappaport BA and Rauschkolb C, 2009. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. Pain, 146(3), pp.238–244. [PubMed: 19836888]

2. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL and Poole RM, 2001. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain, 94(2), pp.149–158. [PubMed: 11690728]

3. Griffith JW, Stephens-Shields AJ, Hou X, Naliboff BD, Pontari M, Edwards TC, Williams DA, Clemens JQ, Afari N, Tu F and Lloyd RB, 2016. Pain and urinary symptoms should not be combined into a single score: psychometric findings from the MAPP Research Network. The Journal of urology, 195(4 Part 1), pp.949–954. [PubMed: 26585679]

4. Naliboff BD, Locke K Jr, Schrepf AD, Griffith JW, Moldwin R, Krieger JN, Rodriguez LV, Clemens JQ, Lai HH, Sutcliffe S and Taple BJ, 2022. Reliability and Validity of Pain and Urinary Symptom Severity Assessment in Urological Chronic Pelvic Pain: A MAPP Network Analysis. The Journal of urology, 207(6), pp.1246–1255. [PubMed: 35060778]

5. Naliboff BD, Stephens AJ, Lai HH, Griffith JW, Clemens JQ, Lutgendorf S, Rodriguez LV, Newcomb C, Sutcliffe S, Guo W and Kusek JW, 2017. Clinical and psychosocial predictors of urological chronic pelvic pain symptom change in 1 year: a prospective study from the MAPP Research Network. The Journal of urology, 198(4), pp.848–857. [PubMed: 28528930]

6. Clemens JQ, Kutch JJ, Mayer EA, Naliboff BD, Rodriguez LV, Klumpp DJ, Schaeffer AJ, Kreder KJ, Clauw DJ, Harte SE and Schrepf AD, Williams DA, Andriole GL, Lai HH, Buchwald D, Lucia MS, van Bokhoven A, Mackey S, Moldwin RM, Pontari MA, Stephens-Shields AJ, Mullins C, Landis JR, 2020. The multidisciplinary approach to the study of chronic pelvic pain (MAPP) research network*: Design and implementation of the Symptom Patterns Study (SPS). Neurourology and Urodynamics, 39(6), pp.1803–1814. [PubMed: 32578257]

7. Freynhagen R, Baron R, Gockel U and Tölle TR, 2006. Pain DETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. Current medical research and opinion, 22(10), pp.1911–1920. [PubMed: 17022849]

8. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Häuser W, Katz RS, Mease P, Russell AS, Russell IJ and Winfield JB, 2011. Fibromyalgia criteria and severity scales for clinical and epidemiological studies: a modification of the ACR Preliminary Diagnostic Criteria for Fibromyalgia. The Journal of rheumatology, 38(6), pp.1113–1122 [PubMed: 21285161]

9. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Häuser W, Katz RL, Mease PJ, Russell AS, Russell IJ and Walitt B, 2016. Revisions to the 2010/2011 fibromyalgia diagnostic criteria. InSeminars in arthritis and rheumatism 2016 Dec 1 (Vol. 46, No. 3, pp. 319–329). [PubMed: 27916278]

10. Lai HH, Newcomb C, Harte S, Appleby D, Ackerman AL, Anger JT, Nickel JC, Gupta P, Rodriguez LV, Landis JR and Clemens JQ, 2021. Comparison of deep phenotyping features of UCPPS with and without Hunner lesion: A MAPP-II Research Network Study. Neurourology and urodynamics, 40(3), pp.810–818. [PubMed: 33604963]

11. Zeger SL, Liang KY and Albert PS, 1988. Models for longitudinal data: a generalized estimating equation approach. Biometrics, pp.1049–1060. [PubMed: 3233245]

12. Gallop RJ, Crits-Christoph P, Muenz LR and Tu XM, 2003. Determination and interpretation of the optimal operating point for ROC curves derived through generalized linear models. Understanding statistics, 2(4), pp.219–242.

13. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N and Carr DB, 2005. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain, 113(1), pp.9–19. [PubMed: 15621359]
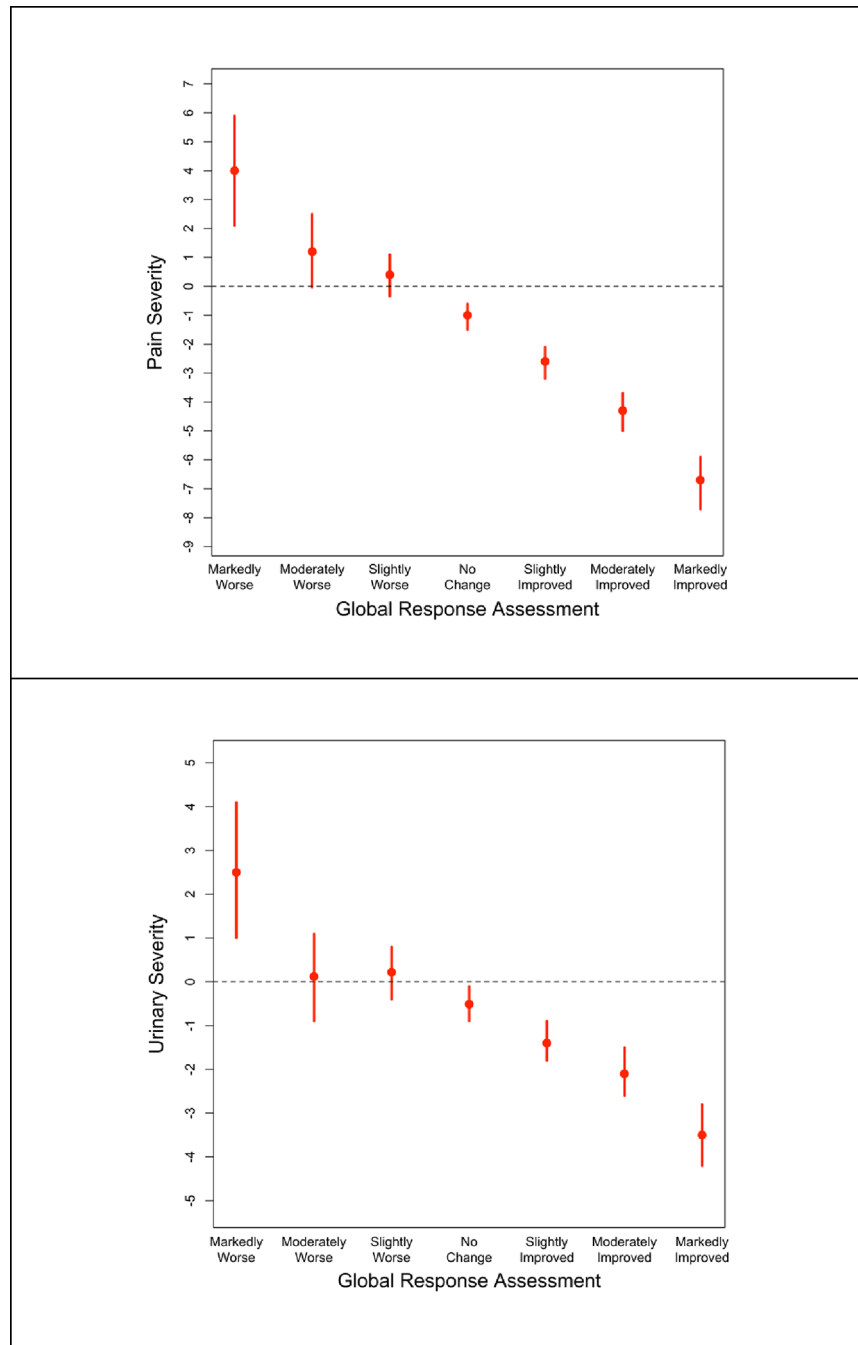
Figure 1.
Regression-based estimates and 95% CIs of absolute clinically important difference for
Pelvic Pain Severity (PPS, top) and Urinary Symptom Severity (USS, bottom)
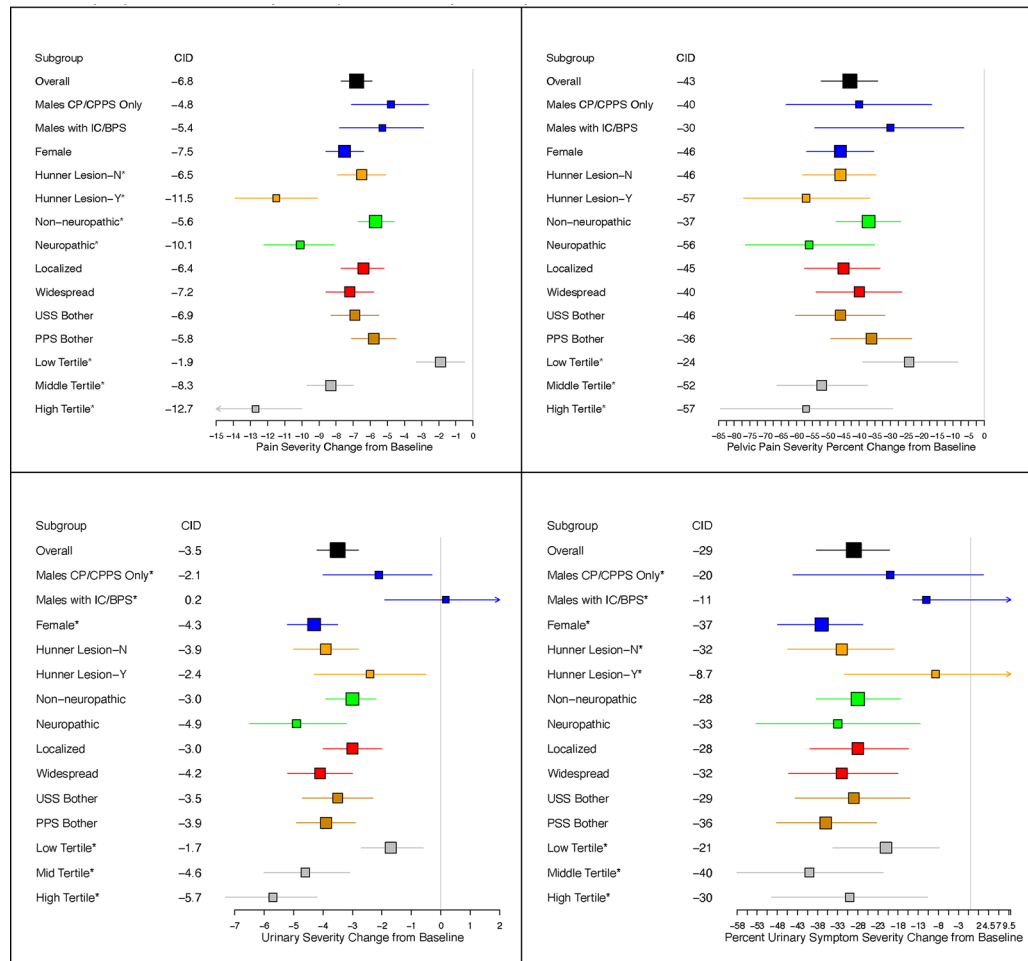
Figure 2.
Estimates of Absolute (left) and Relative (right) Clinically Important Differences for Pelvic Pain Severity (PPS, top) and Urinary Symptom Severity (USS, bottom). Marker size is proportional to sample size; whiskers represent pointwise 95% confidence intervals.

Table 1.

Participant Characteristics

Characteristic			N	Median (25th, 75th percentile) or n (%)
Sex/Diagnosis		Male CP/CPPS Only	537	126 (22)
		Male with IC/BPS		67 (12)
		Female		385 (67)
Age			537	44 (32, 59)
Race		White	537	479 (89)
		Black		25 (5)
		Multiracial		16 (3)
		Other/Unknown		17 (3)
Ethnicity		Hispanic	536	35 (7)
RICE[1] Subtype:		Painful Filling	537	40 (7)
		Painful Urgency		105 (20)
		Both		327 (61)
Cystoscopy Completed			537	363 (68)
Hunner Lesions		No	537	179 (33)
		Yes		28 (5.2)
		Unknown		171 (32)
		No Cystoscopy		159 (30)
Likely Neuropathic		Pain Detect >12	527	186 (35)
Widespreadness		Fibromyalgia Scale >7	537	273 (51)
PPS Bother	Baseline		510	366 (72)
	Month 3		434	290 (67)
	Month 6		459	306 (67)
Pain Severity (0–28)	Baseline		536	16 (12, 19)
	Month 3		469	13 (9, 18)
	Month 6		491	13 (9, 17)
Urinary Severity (0–25)	Baseline		535	13 (7, 17)
	Month 3		468	11 (6, 16
	Month 6		491	11 (6, 16)
BPI-6[3] (0–10)	Baseline		525	4 (3, 5)
	Month 3		458	4 (2, 5)
	Month 6		483	3 (2, 5)
GUPI-4[4] (0–10)	Baseline		536	5 (3, 6)
	Month 3		470	4 (3, 6)
	Month 6		492	4 (2, 5)
Global Response Assessment (GRA)	Month 3	Markedly Worse	474	9 (1.9)
		Moderately Worse		19 (4.0)

Characteristic			N	Median (25th, 75th percentile) or n (%)
		Slightly Worse		56 (12)
		No change		175 (37)
		Slightly Improved		108 (23)
		Moderately Improved		74 (16)
		Markedly Improved		33 (7)
	Month 6	Markedly Worse	492	6 (1.2)
		Moderately Worse		20 (4.0)
		Slightly Worse		51 (10)
		No change		165 (34)
		Slightly Improved		124 (25)
		Moderately Improved		83 (17)
		Markedly Improved		43 (9)

[1.] RAND Interstitial Cystitis Epidemiology Study Criteria

[2.] Complex Multi-Symptom Inventory

[3.] Brief Pain Inventory, Item 6

[4.] Genitourinary Pain Inventory, Item 4

Table 2.

Estimated Spearman Correlation Coefficients at Baseline and 3 Month Follow Up

		PSS	USS	BPI-6	GUPI-4
Baseline	PSS[1]	1.0	-	-	-
	USS[2]	0.53	1.0	-	-
	BPI-6[3]	0.66	0.45	1.0	
	GUPI-4[4]	0.75	0.44	0.73	1.0
Change (3 months)	GRA[5]	0.44	0.28	0.43	0.40
	PSS	-	-	-	-
	USS	0.35	-	-	-
	BPI-6[1]	0.45	0.19	-	-
	GUPI-4[2]	0.74	0.29	0.4	1.0
Percent Change (3 months)	GRA[3]	0.45	0.27	0.46	0.44
	PSS	1.0	-	-	-
	USS	0.32	1.0	-	-
	BPI-6[1]	0.44	0.19	1.0	-
	GUPI-4[2]	0.74	0.26	0.46	1.0

[1] Pelvic Pain Severity (0–28)

[2] Urinary Symptom Severity (0–25)

[3] Brief Pain Inventory, Item 6 (0–10)

[4] Genitourinary Pain Index, Item 4 (0–10)

[5] Global Response Assessment (1–7)

Table 3.

ROC-based Thresholds for Pelvic Pain Severity and Urinary Symptom Severity using 'Markedly Improved' as the Definition of Responder

			Absolute Change		Percent Change		Attained Score	
			CID	AUC	CID	AUC	CID	AUC
PPS	Subgroup	Overall	−4	0.75	−30	0.76	10	0.80
	Sex-Diagnosis	Male CP/CPPS Only	−3	0.69	−25	0.66	9	0.76
		Male with IC/BPS	−2	0.68	−10	0.67	12	0.76
		Female	−5	0.78	−30	0.80	10	0.83
	Hunner Lesions	Yes	−6	0.91	−30	0.91	10	0.87
		No	−4	0.78	−30	0.80	10	0.83
	PainDETECT Type	Likely Neuropathic	−7	0.90	−40	0.90	12	0.86
		Likely Non-neuropathic	−4	0.71	−25	0.70	9	0.75
	Most Bothersome Symptom Type	PPS Bother	−4	0.77	−20	0.78	11	0.83
		USS Bother	−4	0.68	−35	0.70	9	0.72
	Widespreadness	Widespread	−5	0.82	−25	0.82	11	0.80
		Localized	−4	0.69	−30	0.71	8	0.80
	Baseline Tertile	Low tertile	−1	0.57	−10	0.58	7	0.68
		Middle Tertile	−5	0.84	−35	0.85	10	0.86
		High Tertile	−9	0.95	−40	0.94	14	0.93
USS	Subgroup	Overall	−2	0.68	−20	0.71	8	0.69
	Sex-Diagnosis	Male CP/CPPS Only	−2	0.64	−20	0.62	6	0.53
		Male with IC/BPS	0	0.35	7	0.34	10	0.55
		Female	−3	0.76	−25	0.80	8	0.77
	Hunner Lesions	Yes	−1	0.61	−4	0.64	11	0.67
		No	−2	0.69	−20	0.72	8	0.74
	PainDETECT Type	Likely Neuropathic	−3	0.81	−30	0.81	9	0.74
		Likely Non-neuropathic	−2	0.66	−20	0.69	8	0.67
	Most Bothersome Symptom Type	PPS Bother	−3	0.76	−30	0.79	6	0.76
		USS Bother	−2	0.66	−15	0.68	11	0.66
	Widespreadness	Widespread	−3	0.71	−20	0.75	9	0.73
		Localized	−2	0.66	−20	0.68	6	0.65
	Baseline Tertile	Low tertile	−1	0.69	−15	0.68	5	0.59
		Middle Tertile	−4	0.83	−30	0.84	8	0.84
		High Tertile	−3	0.71	−20	0.71	15	0.72