

Published in final edited form as:

*J Urol.* 2013 March ; 189(3): 987–992. doi:10.1016/j.juro.2012.08.257.

## Relationships between Participants' International Prostate Symptom Score and BPH Impact Index Changes and Global Ratings of Change in a Trial of Phytotherapy for Men with Lower Urinary Tract Symptoms

Michael J. Barry, M.D., Alan Cantor, Ph.D., and Claus G. Roehrborn, M.D.

Massachusetts General Hospital (MJB), The University of Alabama at Birmingham (AC), and University of Texas, Southwestern Medical Center (CGR), For the Complementary and Alternative Medicine for Urological Symptoms (CAMUS) Study Group (see Acknowledgements)

### Abstract

**Purpose**—To relate changes in AUA Symptom Index (AUASI) scores with bother measures and global ratings of change among men with lower urinary tract symptoms enrolled in a trial of saw palmetto.

**Materials and Methods**—To be eligible, men were  $\geq 45$  years old, had a peak uroflow  $\geq 4$  ml/sec, and an AUASI score  $\geq 8$  and  $\leq 24$ . Participants self-administered the AUASI, IPSS quality of life item (IPSS QoL), BPH Impact Index (BII) and two global change questions at baseline and 24, 48, and 72 weeks.

**Results**—Among 357 participants, global ratings of “a little better” were associated with mean decreases in AUASI scores from 2.8 to 4.1 points, across three time points. The analogous range for mean decreases in BII scores was 1.0 to 1.7 points, and for the IPSS QoL item 0.5 to 0.8 points. At 72 weeks, for the first global change question, each change measure could discriminate between participants rating themselves at least a little better versus unchanged or worse 70–72% of the time. A multivariable model increased discrimination to 77%. For the second global change question, each change measure correctly discriminated ratings of at least a little better versus unchanged or worse 69–74% of the time, and a multivariable model increased discrimination to 79%.

**Conclusions**—Changes in AUASI scores could discriminate between participants rating themselves at least a little better versus unchanged or worse. Our findings support the practice of powering studies to detect group mean differences in AUASI scores of at least 3 points.

### Keywords

lower urinary tract symptoms; health status index; psychometrics

---

© 2012 American Urological Association. Published by Elsevier Inc. All rights reserved.

Corresponding author: Michael J. Barry, M.D. General Medicine Division, Massachusetts General Hospital, 50 Staniford Street, 9<sup>th</sup> Floor, Boston, MA 02114, Phone 617-726-4106, fax 617-726-4120, mbarry@partners.org.

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

## Introduction

Over the past several years, patient reported outcomes for nonmalignant urologic disease has received increasing attention (1-3). The AUA Symptom Index (AUASI) is widely used in both clinical research and practice as a measure of symptom severity among men with lower urinary tract symptoms (LUTS) attributed to benign prostatic hyperplasia (BPH) (4). The BPH Impact Index (BII) is a higher-level measure of the interference of such lower urinary tract symptoms on patients mental health and activity (5). The International Prostate Symptom Score combines the AUASI with a single separately scored question addressing bother due to lower urinary tract symptoms (IPSS QoL) (6). The BII and the IPSS QoL have been shown to be at least moderately correlated (7)

When the AUASI and BII scores are used as outcome measures, the clinical importance of observed changes in these scores is often judged based on an analysis of data from a large trial of medical therapy for men with LUTS attributed to BPH conducted by the Department of Veterans Affairs Cooperative Studies Program (8). In that study, participants who rated their improvement from baseline to 13 weeks as “slight” had a mean decrease in AUASI score of 3.0 points, and a mean decrease in BII score of 0.5 points. We readdress these thresholds for perceptible differences from baseline among participants in a 72-week randomized dose escalation trial of a saw palmetto extract for men with LUTS attributed to BPH. We also explore the minimum perceptible difference for the IPSS QoL item, which to our knowledge has not been previously addressed. Finally, we analyze the ability of changes in these measures to discriminate between men who rate their urinary symptoms as improved versus unchanged or worse over the course of this study.

## Materials and Methods

The design and main results of the CAMUS trial have been reported previously (9). Briefly, this study was a randomized, placebo-controlled double-blind multicenter trial of increasing doses of saw palmetto fruit extract. Enrollment began in June, 2008 and follow-up was completed in October, 2010. Men were eligible for enrollment if they were  $\geq 45$  years old, had a peak uroflow rate  $\geq 4$  ml/sec, an AUASI score  $\geq 8$  and  $\geq 24$  at two screening visits, and signed informed consent. Participants were recruited at 11 North American sites (see Acknowledgments); the study was approved by their and the Data Coordinating Center's institutional review boards. Participants were randomly assigned equally to receive one, two, and then three 320 mg chocolate-colored gelcaps daily containing a standardized saw palmetto berry extract with dose escalations at 24 (to 640 mg) and 48 weeks (to 960 mg); or an identical number of placebo gelcaps escalated similarly.

Participants self-administered the AUASI at baseline and at each dose escalation at 24 and 48 weeks, as well as at 72 weeks, the end of the study. The AUASI is a self-administered 7 item index assessing frequency of LUTS (range 0-35 points, with higher scores reflecting more frequent symptoms) (4). The change in AUASI score from baseline to 72 weeks was the study's primary outcome. At these same intervals, participants also completed the IPSS quality of life (IPSS QoL) item, “If you were to spend the rest of your life with your urinary condition just the way it is now, how would you feel about that?” The response frame presented seven ordered categories from “*Delighted*” to “*Terrible*” (score range 0-6). Participants also completed the four-item BPH Impact Index (BII), which addresses discomfort, worry, bother, and functional interference from any urinary problems. The BII is scored by addition yielding a range from 0 (no impact) to 13 (great impact). Finally, participants also self-administered two questions addressing their global assessments of change from baseline at 24, 48, and 72 weeks:

Instructions: Circle the appropriate number that best describes your urinary symptoms. Compared to the beginning of the study, how do you feel about your urination now?

1. Worse
2. No change
3. A little better
4. A lot better

Compared to the beginning of the study, how are your urinary symptoms now?

1. Much better
2. Somewhat better
3. A little better
4. About the same
5. A little worse
6. Somewhat worse
7. Much worse

The first question with the asymmetric response frame more closely resembles the global response question from the VA trial originally used to estimate the minimum perceptible difference of the AUASI, while the second question with the symmetric response frame more closely resembles the global response question from the Patient Perception of Study Medication questionnaire (10). All questionnaires were available in English and Spanish.

For this analysis, we examined the relationships between participants' changes in AUASI, IPSS QoL, and BII scores from baseline with their responses to the global assessment of change item at each time point. As there were no significant differences between the saw palmetto extract and placebo group at baseline or at any follow-up point in any of these measures at any time point, we pooled the responses of both groups for the analysis. Moreover, our main paper established that the mean change from baseline in these variables was similar at 24, 48, and 72 weeks; that is, essentially all the improvement was seen by 24 weeks despite subsequent dose escalations (9). We also constructed linear regression models relating mean changes in scores associated with a rating of “a little better” at each follow-up point with participants' baseline scores.

Finally, we constructed logistic regression models to predict participant ratings of at least “a little better,” versus unchanged or worse, on each global response question for each follow-up point. Initially, we used AUASI change, IPSS QOL change, and BII change individually as independent variables to predict participants' global ratings of improvement. Then, we built step-up models using all three variables to maximize model discrimination, as judged by the model's c statistic, using  $P < 0.05$  as the statistical criterion to add an additional independent variable. All analyses were conducted using SAS statistical software version 9.2 (SAS Institute, Cary, North Carolina).

## Results

A total of 369 men were randomized, and 357 participants took at least one dose of study drug and attended at least one follow-up visit. This latter group, the modified intention to treat population, serves as the eligible population for the analyses in this paper. Participants in this modified intention to treat subset had a mean age of about 61 years, and were predominantly well-educated non-Hispanic whites with a mean baseline AUASI score of

14.55 points, IPSS QoL score of 3.21 points, and BII score of 3.55 points. A detailed description of the CAMUS population has been reported previously (9). The numbers of participants who completed an IPSS, BII and the first global change question at 24, 48, and 72 weeks were 316 (88.5%), 332 (93.0%), and 333 (93.3%), respectively. The analogous numbers of participants who completed an IPSS, BII and the second global change questions (these two items were on different questionnaires) were 341 (95.5%), 335 (93.8%), and 333 (93.3%), respectively.

Table 1 presents the correlations of the AUASI scores, IPSS QoL scores, and BII scores at each time point; we had previously reported the baseline correlations of these variables (11). Tables 2,3, and 4 present the mean absolute and percent changes in AUASI scores, IPSS QoL scores, and BII scores for participants according to their ratings on the two questions on global improvement at 24, 48, and 72 weeks, respectively. Because of small numbers, the categories “Somewhat worse” and “Much worse” on the second global response question were pooled. Focusing on participants who rated their global improvement as “a little better” on the first global response question, their mean decrease in AUASI across the three time points ranged from 2.9 to 3.8 points, in IPSS QoL score ranged from 0.5 to 0.7 points, and in BII score ranged from 1.0 to 1.7 points. Thus, the length of the recall period from baseline for the global response question made little difference in the associated mean changes in the three scores. Moreover, at each follow-up point, mean decreases in the three scores for participants who gave the same rating of “a little better” on the second global response question were quite similar.

Figure 1 focuses on participants who rated their global improvement as “a little better” on the first global response question and presents regression lines relating mean absolute changes in AUASI scores, IPSS QoL scores, and BII scores at 24, 48, and 72 weeks, depending on participants' baseline scores. The overall pattern is clear: baseline score has a substantial influence on the changes in scores associated with feeling “a little better” overall. Participants with lower baseline scores who feel “a little better” have relatively small or negligible change scores on each parameter, while participants with higher baseline scores had relatively large mean changes associated with this rating. The findings were consistent for all three measures, as well as across follow-up times.

Finally, focusing on data for the first global change question at 72 weeks, the three change measures individually could correctly discriminate between participants rating themselves at least a little better versus unchanged or worse 70-72% of the time, based on model c statistics (Table 5). The multivariable model increased discrimination to 77%, and all three change scores significantly improved discrimination. For the second global change question at 72 weeks, individual variables correctly discriminated between participants rating themselves at least a little better versus unchanged or worse 69-74% of the time, and the multivariable model increased discrimination to 79%, with only the AUASI and IPSS QoL changes significantly contributing to the higher discrimination. Insights from similar models focusing on 24 and 48 week data were quite similar (data not shown).

## Discussion

In the CAMUS trial, participants' global ratings of “a little better” were associated with mean decreases from baseline in AUASI scores from 2.8 to 4.1 points, across three time points and two ways of asking the global response question. The analogous range for the mean decreases in BII scores was 1.0 to 1.7 points, and for the IPSS QoL item 0.5 to 0.8 points. These estimates of “minimum perceptible differences” in AUASI and BII score changes are slightly larger than the estimates from the VA trial dataset (5), despite CAMUS participants having lower baseline scores, given that in both studies higher baseline scores

were associated with higher minimum perceptible differences. The reasons for these differences are unclear, but in the VA trial data from 13 weeks were used for these estimates, while in the CAMUS study longer recall periods of up to 72 weeks were used. The global change question in the VA trial also used a different response frame of: marked, moderate, slight, no improvement, or worse. The minimum perceptible difference in the IPSS QoL item has not been previously studied, to our knowledge.

A recent analysis of data from the CombAT trial also examined the relationship between changes from baseline in AUASI scores and participants ratings' of global satisfaction with treatment on an item from the Patient Perception of Study Medication questionnaire (10). Satisfaction was rated in seven ordered categories from "very satisfied" to "very dissatisfied." Mean changes in AUASI scores from baseline among participants rating themselves "somewhat satisfied" ranged from -1.41 points among participants with a baseline score of 12 to -11.9 points among participants with a baseline score of 30 (12). Among participants with a baseline score of 16, closest to the mean baseline score among CAMUS participants, a rating of "somewhat satisfied" was associated with a decrease in baseline of -3.7 points, very close to the estimates seen in this study among CAMUS participants rating themselves "a little better."

In the CAMUS study, the AUASI score was significantly correlated with two different measures of bother due to urinary symptoms, and explained about a quarter to a third of the variance in those two bother measures. Moreover, the change in AUASI score from baseline, even after 72 weeks, significantly discriminated between participants who rated themselves at least a little better versus participants who rated themselves unchanged or worse. Additional small but statistically significant increments in discrimination could be achieved by considering changes in the bother measures as well as the AUASI scores. However, the discriminatory abilities of these models are insufficient to replace global response questions as outcome measures if assessing global ratings of response is desired. As new measures of lower urinary tract dysfunction are being contemplated, including for men with symptoms attributed to BPH (<http://www3.niddk.nih.gov/fund/other/MOMUS.pdf>), these results may serve as a comparison.

There are several important limitations to this analysis. First, CAMUS participants had a relatively restricted range of baseline AUASI scores from 8-24 points, which prevented us from fully exploring the relationship between baseline AUASI scores and minimum perceptible changes from baseline. Second, our relatively small sample size and the use of a therapeutic agent that proved to have no greater effect than placebo limits our ability to define with any precision score changes associated with greater degrees of global improvement or deterioration. Finally, comparisons of these results to the results of trials of more effective therapies as in the VA and CombAT trials need to be made cautiously both because the CAMUS study population was less severely symptomatic at baseline, and also received less effective treatment.

## Conclusion

Clinical trials of therapeutic interventions for men with LUTS attributable to BPH are often powered to detect a three point group mean difference in AUASI scores, based on data from the older VA trial (5). In this analysis, participants in the CAMUS trial who rated themselves "a little better" experienced mean AUASI score decreases of 2.8 to 4.1 points, depending on the follow-up time and exactly how the global response question was asked. Therefore, continuing to power trials to detect a group mean difference in AUASI scores of 3 points is unlikely to miss clinically important treatment effects on lower urinary tract symptoms attributed to BPH. However, the effect of baseline scores on the relationship

between global ratings and AUASI score changes suggest that when individuals' symptom responses are being categorized, percentage changes rather than absolute changes in AUASI scores may be the preferable stratifying variable. Changes in AUASI scores from baseline in this clinical trial significantly predicted participants' global ratings of improvement.

## Acknowledgments

### CAMUS Study Group:

Steering Committee Chair

Massachusetts General Hospital, Michael J. Barry, MD

Data Coordinating Center

University of Alabama at Birmingham, O. Dale Williams, PhD (Director), Sreeletha Meleth, PhD (Associate Director), Alan Cantor, PhD

Clinical Sites

New York University, Andrew McCullough, MD (PI through 12/3/10), Christopher Kelly, MD (PI as of 12/3/10), Brianne Goodwin, BSN, RN (Study Coordinator), Laurie Mantor (Study Coordinator), Artrit Butuci (Research Data Associate)

Northern California Kaiser Permanente, Andrew L. Avins, MD, MPH, Harley Goldberg, DO (Co-I), Luisa Hamilton (Study Coordinator), Cynthia Huynh (Research Associate)

Northwestern University Feinberg School of Medicine, Kevin T. McVary, MD (PI), Robert Brannigan, MD (Co-I), Brian Helfand, MD, PhD (Consultant), Maria Velez (Study Coordinator), Nancy Schoenecker, RN, CCRC (Clinical Research Coordinator)

Queens University, J. Curtis Nickel, MD (PI), Alvaro Morales (Co-I), D. Robert Siemens, MD (Co-I), Joe Downey, MSc, CCRP (Study Coordinator), Janet Clark-Pereira, CCRP (Study Coordinator)

University of Colorado Denver, E. David Crawford, MD (PI), Shandra S. Wilson, MD (Co-I), Paul D. Maroni, MD (Co-I), Patricia DeVore, BS (Clinical Research Coordinator), Cliff Jones (Clinical Research Coordinator)

University of Iowa, Karl J. Kreder, MD, MBA (PI), Victoria Sharp, MD, MBA (Co-I), Diane Meyerholz, RN, BSN (Study Coordinator), Mary Eno, RN (Study Coordinator)

University of Maryland, Michael J. Naslund, MD (PI), Ganine Markowitz-Chrystal, MS, CCRC (Study Coordinator)

University of Texas, Southwestern Medical Center, Claus G. Roehrborn, MD (PI), Brad Hornberger, PA-C (Co-I), Allison Beaver, RN (Study Coordinator), Suzie Carter (Data Manager)

Washington University School of Medicine, Gerald L. Andriole, MD (PI), Vivien Gardner, RN, BSN (Study Coordinator), Karen Whitmore (Supervisor Patient Services)

Weill Cornell Medical College, Steven A. Kaplan, MD (PI), Alexis E. Te, MD (Co-I), Noreen Buckley, NP, CCRC (Study Coordinator), Maritza Rodriguez (Medical Assistant)

Yale University School of Medicine, Harris E. Foster, Jr., MD (PI), John W. Colberg, MD (Co-I), Karen Stavris, RN MSN, CCRC (Study Coordinator)

Biostatistics Consultant

University of Arkansas for Medical Sciences, Jeannette Y. Lee, PhD

National Institutes of Health

National Institute of Diabetes, Digestive & Kidney Diseases, John W. Kusek, PhD Leroy M. Nyberg, PhD (through 9/2/09)

National Center for Complementary and Alternative Medicine, Catherine M. Meyers, MD

Office of Dietary Supplements, Joseph M. Betz, PhD

Data Safety Monitoring Board

University of Minnesota VA Medical Center, Timothy J. Wlt, MD, MPH (Chair)

University of Illinois at Chicago, Harry H.S. Fong, Ph.D.

University of Chicago, Glenn S. Gerber, MD

University of Virginia, Mikel Gray, RN, PhD, CUNP, FAAN

HeteroGeneity LLC, Freddie Ann Hoffman, MD

University of North Carolina, Gary Koch, PhD

University of California at Los Angeles, Mark Litwin, MD, MPH

US Environmental Protection Agency, Warren E. Lux, MD

Harvard Medical School, Michael P. O'Leary, MD, MPH

Intercultural Cancer Council, Col (Ret.) James E. Williams, Jr.

Hines VA Hospital Cooperative Studies Program Coordinating Center, Domenic Reda, PhD

**Funding/Support:** This study was funded by cooperative agreements from the National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases: U01 DK63795, U01 DK63797, U01 DK63825, U01 DK63835, U01 DK63866, U01 DK63833, U01 DK63862, U01 DK63840, U01 DK63883, U01 DK63831, U01 DK63778 and U01 DK63788. Support was also provided by the National Center for Complementary and Alternative Medicine and the Office of Dietary Supplements, NIH. Saw palmetto fruit extract and matching placebo was donated by Rottapharm/Madaus, Cologne, Germany. This study was conducted under an Investigational New Drug Application from the Food and Drug Administration. Rottapharm/Madaus had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation or approval of the manuscript. Rottapharm/Madaus provided nonbinding comments to the authors on a draft of the manuscript. NIH scientists representing the funding agencies did participate in the design and conduct of the study as well as review and approval of the manuscript.

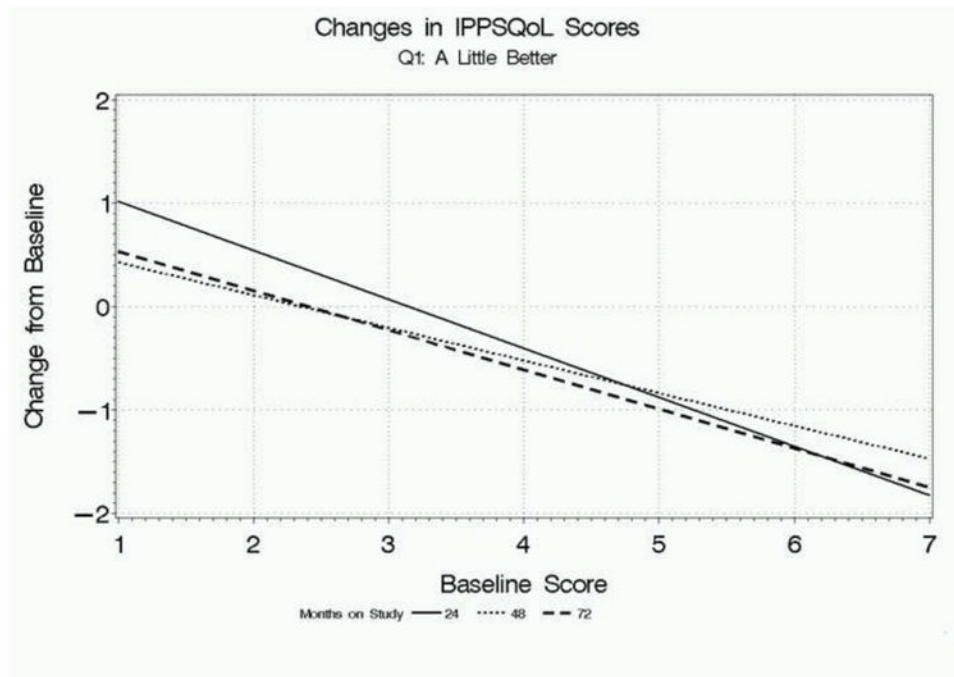
## References

1. Clemens JQ, Markossian TW, Meenan RT, O'Keefe Rosetti MC, Calhoun EA. Overlap of voiding symptoms, storage symptoms, and pain in men and women. *J Urol.* 2007; 178:1354–1358. [PubMed: 17706719]
2. Barry MJ, Link CL, McNaughton Collins MF, McKinlay JB. Overlap of different urologic symptom complexes in a racially and ethnically diverse community-based population of men and women. *Br J Urol Int.* 2007; 101:45–51.
3. 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.* 2010; 64:1260–1278. [PubMed: 20579138]
4. Barry MJ, Fowler FJ, O'Leary MP, et al. *J Urol.* 1992; 148:1549–1557. [PubMed: 1279218]
5. Barry MJ, Fowler PJ, O'Leary MP, et al. Measuring disease-specific health status in men with benign prostatic hyperplasia. *Med Care.* 1995; 4:AS145–155. [PubMed: 7536866]
6. Barry, MJ.; Adolfsson, J.; Batista, JE., et al. Measuring symptoms and health impact of benign prostatic hyperplasia and its treatments. In: Denis, L.; Griffiths, K.; Khoury, S., et al., editors. 4<sup>th</sup> International Consultation on BPH. Plymouth, U.K: Plymbridge Distributors, Ltd.; 1998. p. 265-321.
7. O'Leary MP, Wei JT, Roehrborn CG, et al. *BJU Int.* 2008; 101:1531–1535. [PubMed: 18445080]

8. Barry MJ, Williford WO, Chang Y, et al. BPH-specific health status measures in clinical research: how much change in the AUA symptom index and the BPH impact index is perceptible to patients? *J Urol.* 1995; 154:1770–1774. [PubMed: 7563343]
9. Barry MJ, Meleth S, Lee JY, et al. Effect of increasing doses of saw palmetto extract on lower urinary tract symptoms: a randomized trial. *JAMA.* 2012; 306:1344–1351. [PubMed: 21954478]
10. Barkin J, Roehrborn C, Siami P, et al. Effect of dutasteride, tamsulosin, and the combination on patient-reported quality of life and treatment satisfaction in men with moderate-to-severe benign prostatic hyperplasia: 2-year data from the CombAT trial. *BJU Int.* 2009; 103:919–926. [PubMed: 19239460]
11. Barry MJ, Avins AL, Meleth S, et al. Performance of the American Urological Association Symptom Index with and without an additional urge incontinence item. *Urology.* 2011; 78:550–554. [PubMed: 21741692]
12. Roehrborn CG, Wilson TH, Black LK. Quantifying the contribution of symptom improvement to satisfaction of men with moderate to severe benign prostatic hyperplasia: 4-year data from the CombAT trial. *J Urol.* 2012; 187:1732–1738. [PubMed: 22425127]







**Figure 1.**

- a. Relationship between mean change in AUASI scores from baseline associated with a rating of “a little better” on the first global response question and baseline scores, for each follow-up point.
- b. Relationship between mean change in BII scores from baseline associated with a rating of “a little better” on the first global response question and baseline scores, for each follow-up point.
- c. Relationship between mean change in IPSS QoL scores from baseline associated with a rating of “a little better” on the first global response question and baseline scores, for each follow-up point.

**Table 1**

Spearman correlation coefficients between AUASI, IPSS QoL, and BII scores at 24, 48, and 72 weeks (Ns vary at 24 and 48 weeks due to item nonresponse).

	24 weeks (N=348-349)			48 weeks (N=333-336)			72 weeks (N=357)		
	AUASI score	IPSS QoL Score	BII score	AUASI Score	IPSS QoL score	BII score	AUASI score	IPSS QoL score	BII score
AUASI score	1.0			1.0			1.0		
IPSS QoL score	0.56*	1.0		0.59*	1.0		0.61*	1.0	
BII score	0.48*	0.55*	1.0	0.58*	0.57*	1.0	0.58*	0.60*	1.0

\* P<.0001

**Table 2**

Mean absolute and percent changes with 95% confidence intervals (95%CI) in AUASI scores, IPSS QoL scores, and BII scores from baseline to 24 weeks among CAMUS participants, by their global ratings of improvement.

Global Response Rating	Mean AUASI score change (95%CI)	Mean IPSS QoL change (95%CI)	Mean BII score change (95%CI)
Question #1 (N=316)			
Worse (N=18)	+3.39 (+1.18,+5.60) +24% (+6%,+43%)	+0.89 (+0.33,+1.45) +20% (-2%,+41%)	+0.44 (-0.72,+1.60) +2% (-24%,+29%)
No change (N=178)	-0.74 (-1.48,+0.01) -3% (-9%,+3%)	-0.09 (-0.23,+0.05) +1% (-5%,+7%)	-0.63 (-0.93,-0.32) -10% (-23%,+4%)
A little better (N=91)	-3.49 (-4.63,-2.36) -22% (-30%,-13%)	-0.54 (-0.78,-0.30) -15% (-23%,-6%)	-1.31 (-1.89,-0.73) -22% (-38%,-5%)
A lot better (N=29)	-8.52 (-10.39,-6.65) -51% (-60%,-42%)	-1.62 (-2.15,-1.09) -44% (-55%,-33%)	-2.72 (-3.71,-1.74) -55% (-76%,-33%)
Question #2 (N=341)			
Much better (N=22)	-8.73 (-11.65,-5.81) -55% (-70%,-41%)	-1.91 (-2.56,-1.25) -53% (-66%,-41%)	-3.05 (-4.16,-1.93) -70% (-86%,-54%)
Somewhat better (N=36)	-5.53 (-7.07,-3.99) -35% (-47%,-24%)	-0.86 (-1.21,-0.52) -27% (-37%,-17%)	-1.83 (-2.69,-0.98) -42% (-62%,-22%)
A little better (N=57)	-3.11 (-4.64,-1.57) -16% (-26%,-7%)	-0.46 (-0.72,-0.19) -9% (-18%,+0%)	-1.61 (-2.46,-0.77) -13% (-38%,+12%)
No change (N=197)	-0.99 (-1.66,-0.32) -5% (-11%,+0%)	-0.12 (-0.26,+0.02) +0% (-6%,+5%)	-0.56 (-0.83,-0.29) -12% (-23%,-1%)
A little worse (N=21)	+1.57 (-0.29,+3.43) +12% (-3%,+27%)	+0.19 (-0.30,+0.68) -1% (-12%,+10%)	+0.19 (-0.73,+1.11) +25% (-41%,+92%)
Somewhat/Much Worse (N=8)	+4.88 (+2.90,+6.85) +32% (+17%,+47%)	+1.38 (+0.75,+2.00) +44% (+18%,+69%)	+1.75 (-0.13,+3.63) +55% (-31%,+141%)

**Table 3**

Mean absolute and percent changes with 95% confidence intervals (95%CI) in AUASI scores, IPSS QoL scores, and BII scores from baseline to 48 weeks among CAMUS participants, by their global ratings of improvement.

Global Response Rating	Mean AUASI score change (95%CI)	Mean IPSS QoL change (95%CI)	Mean BII score change (95%CI)
Question #1 (N=332)			
Worse (N=28)	+0.68 (-1.43,+2.79) +8% (-6%,+22%)	+0.29 (-0.17,+0.74) +15% (-1%,+31%)	+0.68 (-0.21,+1.57) +39% (+4%,+73%)
No change (N=180)	-1.56 (-2.23,-0.88) -10% (-15%,-5%)	-0.09 (-0.25, +0.06) +2% (-5%,+10%)	-0.53 (-0.83,-0.23) -8% (-22%,+6%)
A little better (N=95)	-2.86 (-3.96,-1.77) -17% (-25%,-9%)	-0.62 (-0.84,-0.40) -17% (-25%,-10%)	-1.02 (-1.53,-0.52) -13% (-32%,+6%)
A lot better (N=29)	-7.24 (-9.17,-5.31) -49% (-60%,-38%)	-1.62 (-2.14,-1.10) -42% (-52%,-33%)	-3.21 (-4.31,-2.10) -75% (-90%,-61%)
Question #2 (N=335)			
Much better (N=18)	-8.72 (-10.96,-6.49) -59% (-70%,-48%)	-1.83 (-2.60,-1.07) -50% (-62%,-37%)	-3.61 (-5.09,-2.13) -83% (-99%,-67%)
Somewhat better (N=41)	-3.76 (-5.26,-2.25) -27% (-39%,-16%)	-0.93 (-1.22,-0.63) -30%,-40%,-20%)	-1.53 (-2.30,-0.75) -34% (-56%,-11%)
A little better (N=56)	-2.79 (-4.30,-1.27) -14% (-24%,-5%)	-0.48 (-0.76,-0.20) -10% (-18%,-2%)	-1.15 (-1.85,-0.45) -15% (-41%,+11%)
No change (N=184)	-1.60 (-2.28,-0.92) -10% (-15%,-5%)	-0.14 (-0.30,+0.02) +1% (-6%,+9%)	-0.60 (-0.89,-0.30) -13% (-24%,-1%)
A little worse (N=27)	+0.00 (-1.86,+1.86) +2% (-11%,+16%)	-0.04 (-0.41,+0.48) +0.04 (-0.41,+0.48)	+0.41 (-0.46,+1.27) +57% (-1%,+114%)
Somewhat/Much Worse (N=7)	+2.71 (-3.32,+8.75) +20% (-17%,+57%)	+0.71 (-0.17,+1.59) +26% (-1%,+53%)	+2.14 (-0.21,+4.50) +93% (-20%,+206%)

**Table 4**

Mean absolute and percent changes with 95% confidence intervals (95%CI) in AUASI scores, IPSS QoL scores, and BII scores from baseline to 72 weeks among CAMUS participants, by their global ratings of improvement.

Global Response Rating	Mean AUASI score change (95%CI)	Mean IPSS QoL change (95%CI)	Mean BII score change (95%CI)
Question #1 (N=333)			
Worse (N=28)	+1.46 (-0.80,+3.73) +14% (-5%,+33%)	+0.54 (+0.07,+1.00) +25% (+4%,+45%)	+1.11 (-0.05,+2.28) +78% (+6%,+149%)
No change (N=174)	-1.16 (-1.90,-0.41) -7% (-13%,-1%)	-0.06 (-0.21,+0.08) +3% (-4%,+9%)	-0.52 (-0.82,-0.23) -11% (-23%,+1%)
A little better (N=87)	-3.67 (-4.78,-2.55) -24% (-31%,-16%)	-0.70 (-0.92,-0.49) -20% (-27%,-13%)	-1.67 (-2.13,-1.21) -41% (-53%,-29%)
A lot better (N=44)	-8.32 (-9.92,-6.72) -53% (-62%,-44%)	-1.62 (-2.09,-1.14) -42% (-55%,-29%)	-2.73 (-3.54,-1.92) -63% (-78%,-47%)
Question #2 (N=333)			
Much better (N=32)	-9.41 (-11.10,-7.72) -61% (-70%,-52%)	-2.01 (-2.55,-1.46) -53% (-66%,-40%)	-3.16 (-4.14,-2.17) -70% (-88%,-52%)
Somewhat better (N=34)	-4.65 (-6.54,-2.76) -29% (-41%,-17%)	-0.79 (-1.21,-0.38) -23% (-38%,-8%)	-1.65 (-2.54,-0.75) -40% (-64%,-16%)
A little better (N=58)	-4.12 (-5.42,-2.83) -26% (-34%,-18%)	-0.78 (-1.03,-0.52) -21% (-29%,-14%)	-1.63 (-2.16,-1.10) -37% (-51%,-23%)
No change (N=167)	-1.31 (-2.07,-0.55) -8% (-14%,-2%)	-0.10 (-0.24,+0.05) 0.0% (-6%, +6%)	-0.71 (-1.00,-0.42) -17% (-28%,-6%)
A little worse (N=31)	+1.55 (-0.53,+3.63) +13% (-2%,+27%)	+0.52 (+0.04,+0.99) +34% (+6%,+61%)	+0.58 (-0.37,+1.54) +40% (-7%,+86%)
Somewhat/Much worse (N=11)	+1.82 (-0.87,+ 4.50) +15% (-7%,+36%)	+0.55 (+0.19,+0.90) +18% (+5%,+32%)	+2.24 (+0.42,+4.07) +125% (-37%,+287%)

**Table 5**

Logistic regression models predicting participants' global responses of at least a little better versus unchanged or worse on each global response question at 72 weeks.

Question #1 (N=131 at least a little better, 202 unchanged or worse)		
<u>Univariate models</u>	P value	Model C statistic
AUASI score change	<.0001	0.714
IPSS QoL score change	<.0001	0.721
BII score change alone	<.0001	0.698
<u>Stepwise multivariable model</u>		0.771
IPSS QoL score change	<.0001	
AUASI score change	.0023	
BII score change	.0344	
Question #2 (N=124 at least a little better, 209 unchanged or worse)		
<u>Univariate model</u>	P value	Model C statistic
AUASI score change	<.0001	0.738
IPSS QoL score change	<.0001	0.743
BII score change alone	<.0001	0.690
<u>Stepwise multivariable model</u>		0.789
IPSS QoL score change	<.0001	
AUASI score change	<.0001	
BII score change	0.243 (not entered)	