

## ARTICLE



Special Edition on Neuro-urology

# Validity of the Urinary Symptom Questionnaires for people with neurogenic bladder (USQNB) who void or use indwelling catheters

 Rochelle E. Tractenberg<sup>1,2,3✉</sup>, Jamie K. Frost<sup>1</sup>, Futoshi Yumoto<sup>1</sup>, Amanda K. Rounds<sup>4,5</sup>, Inger H. Ljungberg<sup>4,5</sup> and Suzanne L. Groah<sup>3,4</sup>

© The Author(s), under exclusive licence to International Spinal Cord Society 2021

**STUDY DESIGN:** Descriptive Psychometrics Study

**OBJECTIVES:** Neurogenic lower urinary tract dysfunction (NLUTD), or “neurogenic bladder” is a common and disruptive condition for individuals with spinal cord injury (SCI) and disease (including multiple sclerosis, MS). Our team has developed patient-centered instruments of urinary symptoms specific to patients with NLUTD, across bladder management methods. Validity evidence is needed to support the use of two new instruments, Urinary Symptom Questionnaires for people with Neurogenic Bladder (USQNB) for those who manage their bladder with indwelling catheters (IDC), or who void (V).

**SETTING:** Online surveys completed by individuals in the United States with NLUTD due to either SCI or MS who manage their bladder with indwelling catheters (SCI,  $n = 306$ ; MS,  $n = 8$ ), or by voiding (SCI,  $n = 103$ ; MS,  $n = 383$ ). A total of  $n = 381$  USQNB-IDC respondents (five control groups), and 351 USQNB-V respondents (four control groups), contributed to our convergent and divergent validity evidence.

**METHODS:** Data were collected online to estimate key aspects of psychometric validity (*content*, reflection of the construct to be measured; *face*, recognizability of the contents as representing the construct to be measured; *structural*, the extent to which the instrument captures recognizable dimensions of the construct to be measured). *Divergent* and *convergent* validity evidence was derived from multiple control groups, while evidence of *criterion* validity was derived from attribution of each item to their experience “with a UTI”.

**RESULTS:** Evidence of face, content, criterion, convergent, and divergent validity was compiled for each instrument.

**CONCLUSIONS:** The instruments demonstrate adequate, multi-dimensional, validity evidence to recommend their use for decision-making by patients, clinicians, and researchers.

*Spinal Cord* (2021) 59:948–958; <https://doi.org/10.1038/s41393-021-00666-w>

**INTRODUCTION**

Neurogenic lower urinary tract dysfunction (NLUTD), also referred to as neurogenic bladder, is defined as “lower urinary tract dysfunction due to disturbance of the neurological control mechanism” and is a common and disruptive condition for individuals with spinal cord injury and disease (SCI/D) including multiple sclerosis (MS) and spina bifida [1]. Clinically, NLUTD has been classified in different ways, including based on urodynamics, neurological outcomes, or bladder function [2]. Urinary tract infection (UTI) is a common occurrence in individuals with NLUTD and leads to significant additional morbidity and mortality [3–7], beyond that of the NLUTD. Despite the frequency of UTI occurrence and its negative impact, NLUTD-associated UTI definitions and diagnostic criteria have not been standardized

[8–12]. A key criterion of UTI diagnosis is the presence of symptoms, and as such, individuals with NLUTD due to SCI/D have identified urinary symptoms that are not included in, or are explicitly excluded from, UTI diagnostic consideration [13].

Available instruments relevant to bladder function among people with NLUTD have focused on the assessment of *function* [14] and the evaluation of *quality of life* [15, 16], and thus cannot be utilized in identifying UTI or signs and symptoms that may precede UTIs. We have previously reported a clinical trial focused on treating bothersome urinary symptoms [17, 18] for individuals with NLUTD and SCI/D, supporting a shift towards more clinical research into treatments for these specific symptoms. Our research team focused on patient-centered patient reported outcomes featuring the bothersome urinary signs and symptoms

<sup>1</sup>Collaborative for Research on Outcomes and –Metrics, Washington, DC, USA. <sup>2</sup>Departments of Neurology; and Biostatistics, Bioinformatics & Biomathematics, Georgetown University Medical Center, Washington, DC, USA. <sup>3</sup>Department of Rehabilitation Medicine, Georgetown University Medical Center, Washington, DC, USA. <sup>4</sup>MedStar National Rehabilitation Hospital, Washington, DC, USA. <sup>5</sup>MedStar Health Research Health Institute, Hyattsville, MD, USA. ✉email: rochelle.tractenberg@gmail.com

Received: 10 June 2020 Revised: 25 June 2021 Accepted: 28 June 2021

Published online: 4 August 2021

of NLUTD experienced by individuals with SCI/D. Specifically, we created and piloted Urinary Symptom Questionnaires for people with Neurogenic Bladder (USQNB), based on focus groups and interviews with individuals recruited according to bladder management methods. We have previously discussed our reliability evidence for the USQNBs for those using intermittent catheterization (IC) in Tractenberg et al. 2018 [19] together with its validity evidence; and reliability evidence for the USQNB instruments specific to those using indwelling catheter (IDC, inclusive of both indwelling urethral and suprapubic) and voiding (V) in Tractenberg et al., in review [20]. The USQNB instruments are intended to aid in UTI diagnosis for individuals with SCI/D, as well as support research and patient self-management with respect to urinary symptoms potentially attributable to NLUTD-related UTI, depending on bladder management. Throughout this work, we have followed the CONSensus-based Standards for the selection of health Measurement Instruments (COSMIN) [21] to guide our design for prospective testing of these instruments, which were created following a patient-centered patient reported outcome development model we published in 2017 [13]. These standards represent key measurement properties arising from psychometrics and education. One of these properties is *reliability*, defined as either the level of association between equivalent forms of a single assessment or the level of consistency among successive administrations (replications) of the assessment [22]. Mokkink et al. elaborate that “reliability” relates to “(t)he degree to which the measurement is free from measurement error” (p. 108) [21]. The other of these is *validity*, which describes the extent to which evidence and theory support the use/interpretation of the summary of the assessment, often a total score (or subscore) [22]. Here we focus attention on the validity evidence for using the USQNB-IDC and USQNB-V, our two newest instruments, continuing our objectives to promote and support research into new interventions for UTI and bothersome urinary signs and symptoms in the presence of NLUTD.

Formal (psychometrically-defined) validity is a multi-dimensional continuum that comprises the evidence supporting the use of an assessment for a specific (decision-making) purpose; this has been extensively discussed and studied for the cognitive and achievement assessment of students/in educational settings [23]. Geisinger discusses the modern perception of validity as “...a unitary concept for which there are five types of evidence to gather to justify the use of a particular measure (p. 631).” [23] These five types include content; relations with other variables of interest; internal structure of the instrument; what the respondent (must be doing) as they answer; and the consequences of the use of the instrument. These features, representing the technical definitions of validity evidence from the field of psychometrics, are consistent with—and go beyond—the COSMIN evaluation approach outlined in Prinsen et al. 2018 (Phase B, step 5 (content validity) and step 7 (criterion validity)) [24].

In this paper we explore the psychometric validity of patient responses on the two newest USQNBs, for indwelling catheter users and voiders, using Geisinger’s set of evidence types<sup>23</sup> to ensure we meet and exceed the COSMIN standards [21, 24]. While there is no ‘gold standard’ for assessing urinary symptoms attributable to UTI, this manuscript describes evidence of the validity of the two newest USQNBs for use by clinicians, researchers, and by individuals to self-monitor and self-manage these symptoms.

## METHODS

All USQNBs were developed following our model for patient-centered patient reported outcomes [13], separately to elicit patient-centered reports of urinary symptoms by individuals with NLUTD according to bladder management method. Approval for the studies was received from the MedStar National Rehabilitation Hospital Institutional Review Board

(V: IRB# 2016-212, IDC: IRB# 2016-088). As discussed below, COSMIN definitions of “validity” were augmented with other standards [22] to ensure that decisions made by clinicians, researchers, and patients would be supported by the items on these instruments. That is, potential users of these instruments need evidence that they will generate information about the respondent’s experiences that is appropriate for the uses to which it will be put [25]. To ensure content validity as well as convergent and divergent validity, we recruited responses from national samples within and outside of our targeted respondents specifically for the purposes of establishing the validity (this paper) and reliability (Tractenberg et al., in review [20]) of these two instruments. Our analyses focused on endorsement of the items, because the impact, frequency, and severity ratings increase the difficulty of generalizing results.

## Recruitment of target participants

As we have reported elsewhere for this sample [20], participants in the patient groups these instruments target were recruited via direct email, social media, and with the assistance of advocacy organizations. Participants were recruited in the United States by English-language advertising through Facebook, via email, and with the assistance of the national (U.S.) advocacy networks in spinal cord injury and multiple sclerosis. All of these outreach and recruitment efforts were advertisements seeking respondents with NLUTD who use the specified bladder management method to visit the URL we established for data collection. “Voiding” was defined as “primarily voids to empty their bladder; that individual may intermittently (no more than once per day) use an intermittent catheter to empty his/her bladder. “Indwelling catheter” was defined to include suprapubic and indwelling urethral types. These were essentially the inclusion criteria for our target groups.

These recruitment efforts targeted individuals with SCI/D who have NLUTD or neurogenic bladder. Partner advocacy groups announced study recruitment. Participants were recruited using a SurveyMonkey link sent as a direct e-mail or posted on our website and Facebook. No identifying information was gathered to be able to identify or re-contact those who completed the survey. We included questions in the survey about time since diagnosis (of SCI or MS) and whether respondents had received a diagnosis of neurogenic bladder or NLUTD or if they did not know if they had this diagnosis. We also asked about respondent experience of UTI (“how many urinary tract infections (UTIs) have you been diagnosed with in your lifetime?”) -but we did not ask for details about any UTI diagnosis. The responses to these ancillary questions were used to ensure we obtained responses from our target population (SCI/D with NLUTD). We then sought responses from participants in our validation (divergent/convergent) populations, described below. No personal or identifying data were collected from any respondent, and initiating responses on the survey was deemed sufficient consent by the IRB.

## Divergent validity participant recruitment

As with the USQNB-IC [19], we sought divergent validity data by having people who did not meet our inclusion criteria—i.e., having no experience with UTI, NLUTD, or both—complete the instruments. Individuals with NLUTD who self-report a history of no UTIs (“how many UTIs have you been diagnosed with in your lifetime?”= none), individuals with chronic mobility impairments (but without SCI or MS) and without NLUTD, and individuals with no mobility impairment, no NLUTD, and no history of UTIs were recruited largely from an inner-city rehabilitation hospital to complete the USQNB-V and USQNB-IDC, according to their bladder management method, if relevant. We sought a minimum of 30 respondents in each of our “divergent validity evidence” groups, planning to close the survey once we reached 50 in any of these groups. We planned a 6-month data collection period, and over time we were able to determine whether we were on track to reach our minimum group size. We determined that, if we could not get even 30 respondents in a subgroup, we would explore collapsing smaller subgroups, e.g., we started with 0 UTIs, 1-2 UTIs, 3-5 UTIs, and >5 UTIs as subgroups for voiders; however, we anticipated these might be too rare to allow us to reach our minimum response targets for each divergent validity group. Thus, we planned to collapse over the UTI-count subgroups, as needed (but not over groups based on other characteristics, like mobility or NLUTD status).

All respondents were asked about each symptom they endorsed within the past year, with instructions, “when considering a symptom, you should first identify what is a normal experience for you. After identifying your normal experience, report any CHANGE from that. If you report “Yes”, you

will then be asked follow-up questions regarding that particular symptom". For symptoms that were endorsed, respondents were asked to describe whether they attributed that symptom "to a urinary tract infection (UTI)". We did not anticipate that "attributed to a UTI" would be consistently interpretable [19] given the facts that symptoms may have occurred more than one time over the year; and that there exists a wide range of diagnostic criteria for "UTI". Although we included attribution options of "always", "sometimes", "rarely", and "never, we planned to only use the response "never attributed to a UTI" to characterize the tendency to experience both the specific symptom *and* a UTI. Thus, contributing to our evidence of convergent and divergent validity, we computed the % of each responding group who *did endorse* a symptom *and never* attributed it to a UTI. Convergent validity evidence comes from similar groups having similar response patterns to those in the target group, while divergent validity evidence comes from dissimilar groups having different response patterns from the target group's patterns. Both of these represent Geisinger's "evidence based on relations with other variables", with additional convergent validity evidence coming from *lower* rates of endorsement of a symptom with never attributing it to a UTI within our target group and validity subgroups with similar UTI lifetime experience, as compared to those with lower levels of UTI experience.

**Materials**

For each instrument (and following the same method as was reported in Tractenberg et al. [19]), bladder management-specific focus groups and

individual interviews were convened to discuss the items that represent the patient experience of urinary symptoms associated with NLUTD. These were reviewed and revised iteratively by clinical subject matter experts, patient experts, and our research team, leading to the instruments reported in this paper.

Once each instrument was finalized, national samples of people with NLUTD who manage their bladders with indwelling catheters (IDC) or by voiding (V) were recruited to complete the new (relevant) instrument (USQNB-IDC or USQNB-V) online, using SurveyMonkey. Participants anonymously filled out one instrument—specific to their management type, by following a URL to the survey site. Each item was presented in English as a query about whether the respondent had experienced it during the past year (yes/no). These instruments contain 26 different items each (plus one item, "other"), and all items are phrased such that endorsement indicates either a greater level than is normal, or more intense experience of something than is usual (e.g., darker urine than normal).

For all instruments, the symptoms have been categorized into one of four different symptom types: those that are clinically actionable (A, "actionable"); those that represent bladder-specific signs and symptoms (B1, "bladder actionable"); those representing characteristics of urine (B2, "urine quality"), and all of the other items (C, "other") that were retained from the initial patient focus groups and interviews, following evaluation by clinicians and investigators. The instruments and classifications are the subject of a new paper currently in preparation; the symptoms appear in Tables 3 and 4, sorted according to the classifications<sup>1,2</sup> shown in Fig. 1A, B in the Results section.

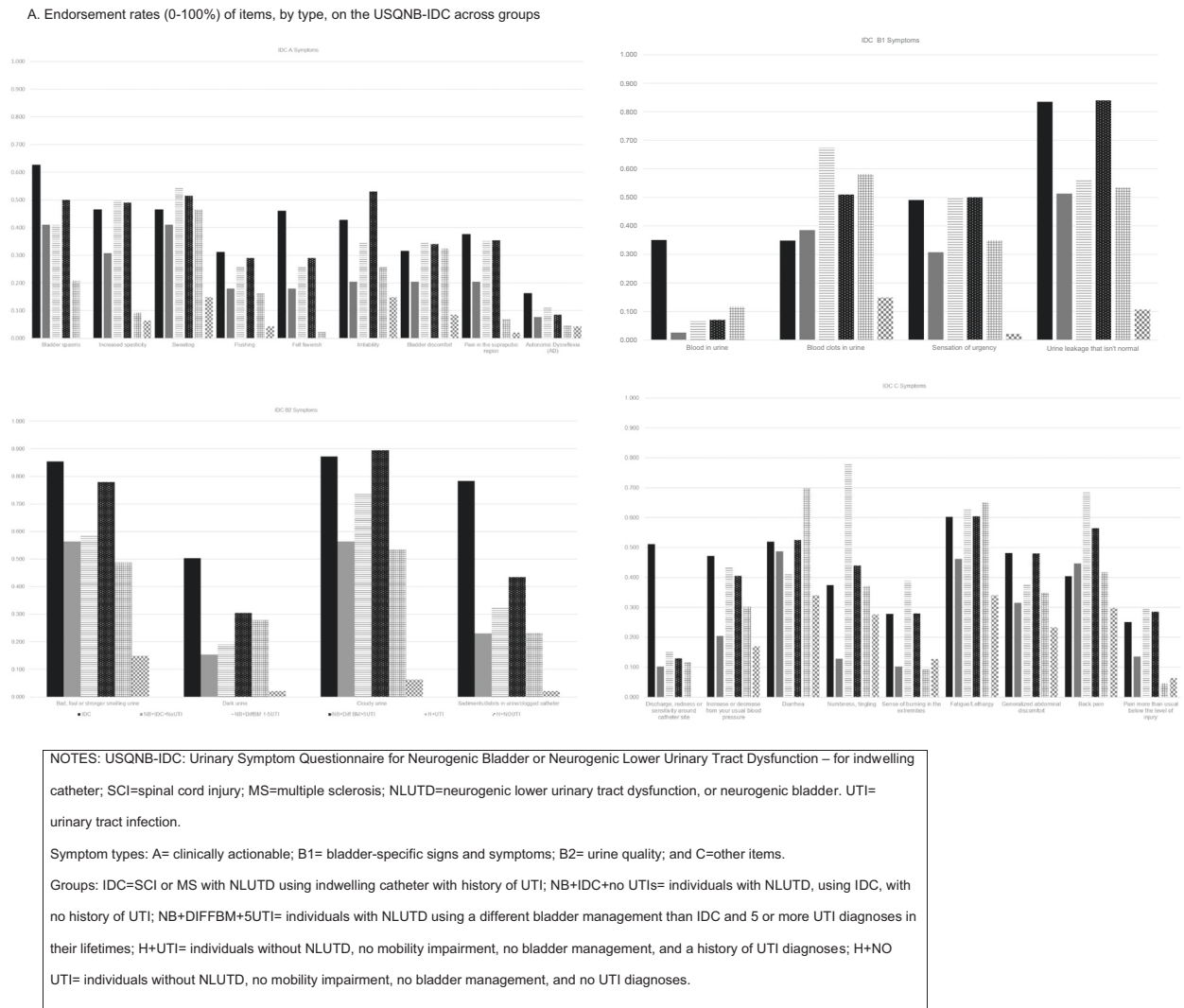
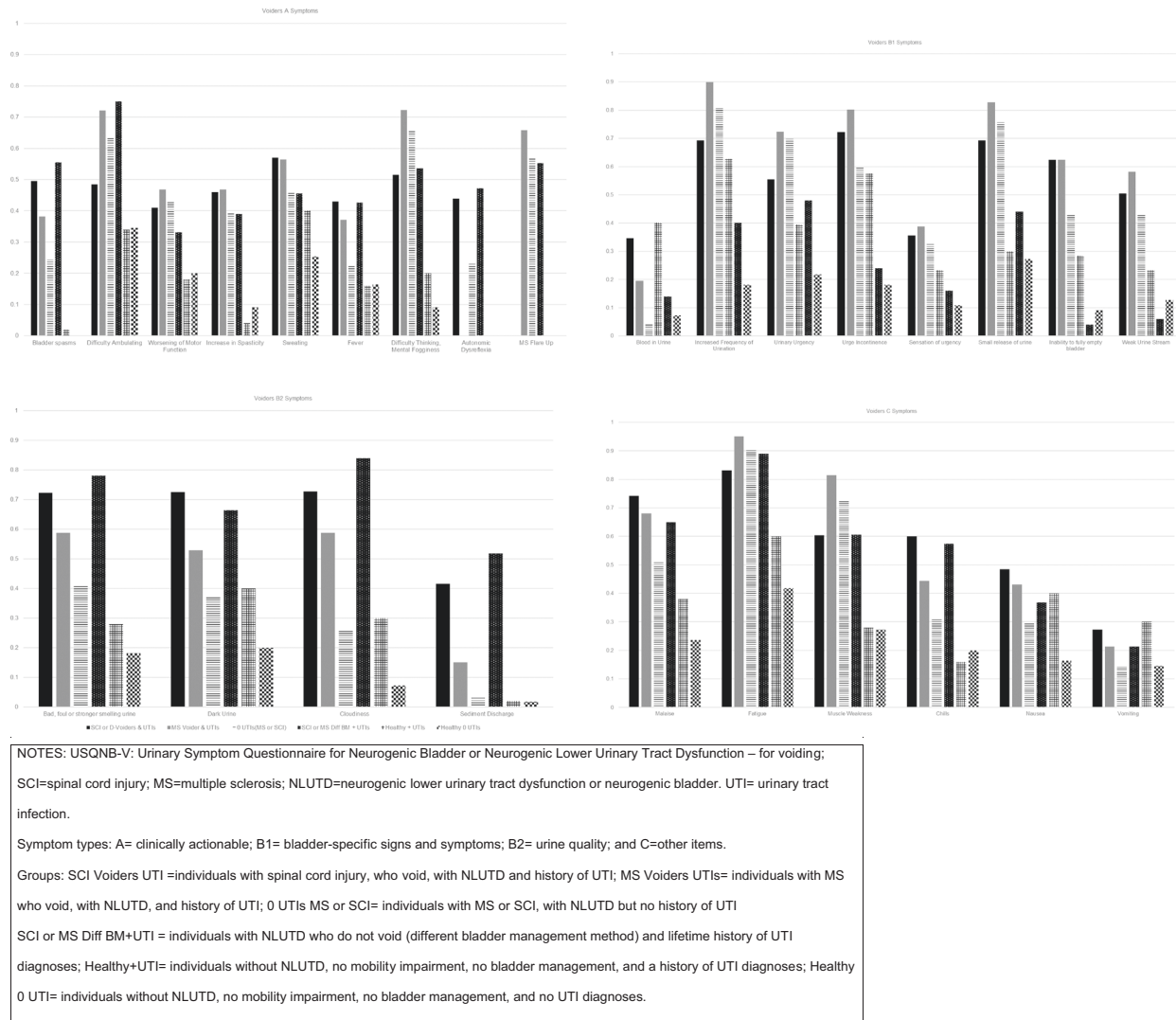


Fig. 1 (Continued).

B. Endorsement of items (0-100%), by type, on the USQNB-V across groups



**Fig. 1** Percentage of those in each responding group endorsing items on the USQNB-IC and USQNB-V. **A** Endorsement of items on the USQNB-IC across groups. **B** Endorsement of items on the USQNB-V across groups.

The analyses described below for each instrument were focused on the endorsement (yes/no) of the items on each instrument. Tables 3 and 4 show the symptom that individuals were asked to endorse if it had been experienced within the previous 12 months. All responses were given “for the past year”, even though we intend for the instruments to be used in a much shorter time frame. This yearlong time frame was featured to ensure that the preliminary validity evidence we collected from our national sample on each instrument was as inclusive as possible.

### Validity evidence

These analyses were planned to generate evidence of instrument validity. We defined *content* validity as the reflection of the construct to be measured—urinary symptoms potentially attributable to UTI for NLUTD depending on bladder management; *face* validity represented by the recognizability of the items as representing the construct to be measured—symptoms that were generated by subject matter experts and similar patients; internal or *structural* validity, or the extent to which the instrument captures recognizable dimensions of the construct to be measured. However, because there is no “gold standard” for urinary symptoms of UTI in this population, we were unable to generate evidence of *criterion* validity, defined by association with a gold standard. Instead, we used the patterns of “never attributed to a UTI” among the target

respondent groups, so that *lower levels* of never attributing the symptom with a UTI would suggest association with UTI.

In addition, we collected evidence on *convergent* (people around the country with similar bladder management and a history of urinary symptoms should experience the symptoms our focus groups identified) and *divergent* (people around the country with similar bladder management and a history of urinary symptoms should NOT experience, or should experience much less frequently, the symptoms our focus groups identified) validity. Divergent validity data came from individuals with NLUTD who do not have a history of UTIs, individuals with chronic mobility impairments (no SCI/D) and without NLUTD, and those with no mobility impairment, no NLUTD, and no history of UTIs. COSMIN criteria [21] define construct and criterion validity as *including* divergent and convergent validity; including both divergent and convergent validity data conforms to Geisinger’s “relations with other variables of interest”. We also used the patterns of “never attributed to a UTI” among the divergent validity respondent groups, so that *higher levels* of never attributing the symptom with a UTI would suggest an important *lack of* association with UTI. Given the challenges in diagnosing UTI arising from variable criteria and the reliance on self-report, the clearest criterion signal comes from the “never attributed” response.

As we argued in our exploration of the psychometric properties of the USQNB-IC [19], structural validity was not expected to yield interpretable



**Table 1.** COSMIN validity criteria and how they were assessed in this study.

COSMIN construct	Definition	Approach/analysis	Comment
Validity-content	Degree to which instrument measures the construct it targets	By development & design (specifically created to achieve this purpose)	Previously published <sup>13</sup> ; detailed descriptive statistics, this paper
Validity – face	Degree to which items “look” as if they are an adequate reflection of the target construct	By development & design, iteratively eliciting and obtaining input from patients and clinicians.	Previously published <sup>13</sup> ; detailed descriptive statistics, this paper
Validity-construct	Degree to which the scores are consistent with expected similarities (convergent) and differences (divergent) between groups	Divergent and convergent validation samples; comparisons of endorsement rates.	Convergent validity: similar endorsement rates across NB groups; divergent validity: endorsement rates for people without NB similar to each other, dissimilar to NB groups.
Validity-criterion	Degree to which the scores reflect a “gold standard”	See construct validity; also, by never attributing of each item to “having a UTI”	No diagnostic gold standard; we use convergent and divergent validity data instead.
Validity-structural	Degree to which the scores are an adequate reflection of the dimensionality of the target construct	Bayesian Network to uncover associated signs and symptoms; Principal axis factoring. <sup>22</sup>	Patient-centered instruments do not have pre-specified dimensionality. Structural information published elsewhere. <sup>22</sup>

NOTES: COSMIN = Consensus-based Standards for the selection of health Measurement Instruments<sup>19</sup>; UTI = urinary tract infection; NB = neurogenic bladder or neurogenic lower urinary tract dysfunction.

results for either of the two new instruments, because this analytic approach to validity is not consistent with either our approach to developing the patient-centered patient reported outcomes, nor with the one-year time frame of our national sample surveys. However, as we have done before, we confirmed this suspicion with exploratory common factor analysis (using principal axis factoring) to model shared covariance among items, and inferred causal model with Bayesian networks, to model the shared Shannon Information [20].

Table 1 below presents the COSMIN validity criteria together with the methods by which evidence of these properties for the instruments was obtained.

### Data analysis

Statistical analyses to generate the validity evidence described above/ listed in Table 1 were carried out using SPSS v. 24 for descriptive statistics; these analyses were based on the yes/no endorsement of all items for the samples by instrument. We identified an endorsement rate minimum cut-off of 10% [19], below which we would want to reconsider the item's inclusion due to low national endorsement over the course of a year, suggesting lower-than-acceptable content or face validity. We previously reported the results of similar analyses of the USQNB-IC [19], and reliability evidence for the USQNB-IDC and USQNB-V [20].

## RESULTS

### Demographics

We checked to ensure respondents in each group (target, convergent, divergent) met the criteria for that group and bladder management method. All respondents on the IDC were included with SCI ( $n = 306$ ) or MS ( $n = 8$ ). All respondents on the USQNB-V with SCI who met criteria for NLUTD, although the survey item asked specifically about *neurogenic bladder* (which might be more familiar to respondents), and were included as “voiders with NLUTD” ( $n = 103$ ). Of the 447 individuals with MS who were recruited to complete the voider instrument, those with MS who indicated they did not have/did not know if they had NB and who did *not* endorse any bladder- or urine-specific symptoms ( $n = 4$ ) were excluded because their NLUTD status was ambiguous, leaving 405 MS voiders with NLUTD in the sample. four people indicated they had a MS diagnosis in the demographics questions, but in another question (“MS flare up”), responded that they did *not* have MS; another person stopped responding after completing fewer than half of the USQNB-V. Of the remaining 442 MS voider responders, 33 indicated that they did not have neurogenic bladder (NB), but instead, indicated they had “some other bladder problem”; and so were excluded from consideration. Of the remaining 409, 185 indicated that they had NB, and 224 reported that they did not have, or did not know if they had, a diagnosis of NB. item was Of

these, 220 individuals were included as “voiders with NB” because they did have a diagnosis of MS or SCI, and they indicated that they *did* experience at least one bladder- or urine-specific (B1 or B2) symptoms on the USQNB-V (thus, endorsing a NLUTD-causing SCI/D condition plus some experience of symptoms associated with NLUTD).

Table 2A, B show the descriptive statistics for our target groups (IDC or V with NB and the specific bladder management), as well as the divergent validity groups. Two individuals in one voider validity group indicated that they had no history of UTI, but when asked if they attributed any USQNB-V symptom that they endorsed to a UTI, one responded “always” on all endorsed items and the other responded “occasionally” on most endorsed items. Their responses (on all items) were excluded from all analyses.

The divergent validity groups included people with MS or SCI who managed their bladders with catheters and who reported a history of UTIs in the following three categories in their lifetime: No UTIs in their lifetime, 1–5 or >5 UTIs. This information was self-reported by respondents, we were unable to confirm category membership. We were also forced to combine the 1–5 and >5 UTIs voider groups into one “history of UTI” group in order to reach our target of 50 (or more) in this group. For the IDC instrument we were able to find individuals in sufficient numbers (at least 50) who self-identified as having lifetime histories of 1–5 UTIs and >5 UTIs, but this was not possible during our recruitment period for the voiders divergent validity groups.

### Validity evidence

Tables 3 and 4 present the percentages of respondents endorsing the symptom within the past year in the target and our convergent/divergent validity groups. Table 3 describes the IDC endorsements across all groups, while Table 4 describes endorsement for Voiders.

The endorsement rates for each item on both instruments were all >20% for our target respondents. The levels of endorsement by the target groups support claims of face validity for both instruments for the target population; all items were endorsed in the national sample. These patterns are clearer in Fig. 1A, B, reflecting the endorsement rates for all groups on the instrument for IDC (1 A) and V (1B), below.

Fig. 1A, B show that the target groups endorse all items on their respective instruments sufficiently to suggest that the items on the instrument have face and content validity; people with history of UTI endorsed items at a greater frequency than those with a history of no UTI on both instruments.

**Table 2. A.** Descriptive statistics, indwelling catheterization (IDC) target and validity groups. B. Descriptive statistics—voiders target (MS, SCI) and validity groups.

Bladder function: bladder management: UTI status: group:	NLUTD IDC > 2 UTIs (target)	NLUTD IDC No UTI (div)	NLUTD No IDC 1-5 UTIs (conv)	NLUTD No IDC > 5 UTIs (conv)	No NLUTD, No IDC No UTIs (div)	No NLUTD, No IDC No UTIs (conv)
# of participants	314 (MS = 8)	40	44	206	51	50
Average age of participants ± SD (range)	54.3 ± 13.2 (8-80)	51.1 ± 19.5 (20-88)	51.6 ± 13.9 (20-75)	50.8 ± 15.4 (4-77)	46.5 ± 16.6 (23-78)	52.0 ± 17.8 (19-85)
Average time living with injury ± SD (range)	19.7 ± 14.6 (1-64)	13.2 ± 12.0 (0-40)	13.8 ± 15.4 (0-67)	21.1 ± 15.8 (1-68)	NA	NA
% of women	38.5	33.3	40.9	40.0	41.1	84.0
Bladder function: Injury/disease: Bladder management: UTI status: group:	NLUTD SCI Void (target)	NLUTD MS Void (target)	NLUTDMS/SCI Void No UTIs (div)	NLUTD MS/SCI No Void <sup>a</sup> + UTIs (conv)	Normal N/A Void UTIs (div)	Normal N/A Void UTIs (conv)
# of participants	104	405	118	137	55	50
Average age of participants ± SD (range)	55.5 ± 13.7 (20-82)	52.8 ± 12.6 (10-80)	51.4 ± 13.3 (23-84)	55.2 ± 12.8 (20-80)	52.2 ± 15.9 (26-82)	52.3 ± 15.9 (19-84)
Average time living with injury ± SD (range)	16.9 ± 15.9 (1-65)	14.9 ± 10.3 (0-53)	12.3 ± 10.2 (0-65)	20.3 ± 10.8 (1-54)		
% of women	50.9	89.1	45.8	50.4	40.0	86.0

NOTES: SCI = spinal cord injury; MS = multiple sclerosis; NLUTD = neurogenic lower urinary tract dysfunction or neurogenic bladder; UTI = urinary tract infection (self-reported lifetime history). Target = patient group targeted by the instrument; div = divergent validity group; conv = convergent validity group. SD = standard deviation. # UTIs = self report of (#) of UTI diagnoses in lifetime; "no UTIs" = self report of no UTI diagnoses in lifetime.

<sup>a</sup> Respondents in this category used some method of bladder management that was "not voiding". SD = standard deviation. "+ UTIs" = self report of at least one UTI diagnosis; "no UTIs" = self report of no UTI diagnoses in lifetime.

Our final analysis of divergent and convergent validity was an examination of association of the endorsed symptoms "to a UTI". As noted, the figures show the percentage of respondents in each group that did endorse each item, but *never attributed it to a UTI* (whether or not they reported ever having had a UTI before). Fig. 2A shows results for IDC, and 2B shows them for V.

Values shown in Fig. 2A, B falling below 50% indicate that the specific symptom was attributed to a UTI (however "UTI" might have been determined by the respondent) at least 50% of the time. Respondents in divergent validity groups without history of UTI were expected to choose "never attributed to a UTI" 100% of the time, and this was observed for both the IDC and V samples. The figures suggest that, while prevalent, many clinically actionable (A) and "other" (C) type symptoms are not typically attributed to having a UTI when they are experienced by our target groups. For the IDC target group, both bladder (B1) and urine quality (B2) type symptoms are usually attributed to UTIs, while for voiders this is true for most urine quality (B2) but not all bladder (B1) type symptoms. In the divergent validity groups (non-target respondents), the rates tended to indicate less attribution to UTI if these were in their inclusion criteria/history, as compared to the target groups; the responding groups we recruited with histories of UTI naturally tended to have higher rates of attribution to UTI (i.e., lower points on these plots showing the percentage in each group that *never attributed* each symptom to a UTI).

**DISCUSSION**

The purpose of this study was to describe the psychometric validity evidence supporting implementation of two new patient-centered patient reported outcome instruments for urinary signs and symptoms in people with NLUTD and who use either indwelling catheterization to manage their bladder (USQNB-IDC), or who void (USQNB-V). Validity is a continuum, and with so many sources it is impossible to coherently summarize all the evidence quantitatively. Evidence of content and face validity comes from the instrument design process and the endorsement rates by national samples. Criterion validity is challenging to document because there is no accepted gold standard for the patients' lived experience, but we were able to estimate this using respondent attribution of each symptom to experience with UTIs. Convergent and divergent validity evidence were obtained both from diverse respondent groups, and from the "never attributed to UTI" item on the survey we administered. Specifically, attribution of USQNB items to UTI tended to be greater for those with greater experience with UTIs, resulting in lower rates of "never attributed to a UTI" as well as higher endorsement rates when compared to respondents with no history of UTI. Convergence across those with similar UTI and bladder management experience was observed in both the endorsement and the attribution, while divergent validity evidence is similarly presented for those without, or with lower levels, of UTI experience and with bladder management that does not involve catheters. Structural validity evidence was found for these instruments with the target groups, and has been reported elsewhere [20], which is relevant despite the lack of a measurement (causal) model in the instrument's development. Our validity evidence is summarized below:

**Face, content, convergence validity evidence**

All items—which were generated by patients themselves in our focus groups, and then integrated with clinician and researcher input—were recognized and endorsed by at least 20% of the national samples using each bladder management method (most were endorsed by at least 50%). We also observed that endorsement rates were highest for all groups who reported a history of UTIs, as compared to corresponding groups reporting

**Table 3.** USQNB-IDC endorsement rates by group (IDC, and convergent/divergent groups).

Symptoms	NLUTD IDC >2 UTIs	NLUTD IDC No UTI	NLUTD No IDC 1–5 UTIs	NLUTD No IDC >5 UTIs	No NLUTD, No UTIs	No NLUTD, + UTIs
n	314	40	51	206	51	50
A1: Bladder spasms	0.626	0.400	0.412	0.507	0.039	0.260
A2: Increased spasticity (not bladder spasms)	0.464	0.300	0.510	0.500	0.059	0.140
A3: Sweating	0.464	0.400	0.529	0.510	0.157	0.500
A4: Flushing of the chest, neck or face	0.312	0.150	0.235	0.294	0.078	0.160
A5: Autonomic Dysreflexia (AD)	0.460	0.175	0.235	0.289	0.000	0.020
A6: Felt feverish	0.427	0.200	0.353	0.525	0.176	0.300
A7: Irritability, mental slowing/ confusion	0.315	0.200	0.392	0.348	0.098	0.380
A8: Bladder discomfort	0.375	0.200	0.360	0.368	0.059	0.140
A9: Pain in the suprapubic region	0.163	0.050	0.120	0.093	0.039	0.060
B1.1: Blood in urine	0.503	0.150	0.196	0.301	0.039	0.300
B1.2: Blood clots in urine	0.350	0.000	0.059	0.068	0.000	0.120
B1.3: Sensation of urinary urgency	0.348	0.375	0.706	0.525	0.196	0.620
B1.4: Urine leakage that isn't normal	0.490	0.300	0.549	0.505	0.039	0.400
B2.1: Bad-smelling, stronger of fouler urine	0.835	0.500	0.608	0.830	0.137	0.540
B2.2: Dark urine	0.854	0.550	0.608	0.777	0.176	0.500
B2.3: Cloudy urine	0.873	0.550	0.745	0.883	0.098	0.540
B2.4: Sediments/debris in urine/ clogged catheter	0.783	0.250	0.294	0.422	0.020	0.240
C1: Diarrhea	0.519	0.500	0.451	0.520	0.353	0.700
C2: Fatigue/Lethargy	0.603	0.450	0.647	0.608	0.373	0.620
C3: Back pain	0.405	0.436	0.700	0.564	0.314	0.420
C4: Numbness, tingling in your extremities	0.375	0.125	0.765	0.446	0.255	0.380
C5: Pain more than usual below the level of injury	0.251	0.128	0.306	0.289	0.059	0.060
C6: Discharge, redness or sensitivity around catheter site	0.511	0.100	0.137	0.123	0.000	0.120
C7: Generalized abdominal discomfort or pain: bloating, pressure or cramping	0.482	0.333	0.400	0.485	0.255	0.340
C8: Sense of burning in the extremities or throughout the body	0.278	0.100	0.392	0.279	0.118	0.100
C9: Increase or decrease from your usual blood pressure	0.472	0.200	0.412	0.402	0.196	0.340

NOTES: Values in the table are rates, so 0.626 = 62.6% of the sample endorsed that item. USQNB-IDC: Urinary Symptom Questionnaire for Neurogenic Bladder or Neurogenic Lower Urinary Tract Dysfunction – for indwelling catheter; SCI = spinal cord injury; MS = multiple sclerosis; NLUTD = neurogenic lower urinary tract dysfunction, or neurogenic bladder. UTI = urinary tract infection. “no UTI” = self report of no UTI diagnoses in lifetime. A = Clinically Actionable symptoms; B1 = Bladder Specific symptoms; B2 = Urine Quality symptoms; C = Constitutional symptoms.

no UTIs. Further, the symptoms on the USQNBs are typical of individuals with NLUTD who experience UTIs, but it is not clear whether the USQNB items are solely specific to NLUTD-related UTI or that they might extrapolate to other populations. Our validity evidence suggests that the items are generally more commonly endorsed by individuals with experience of UTIs than those without a history of UTIs.

#### Divergent validity

Few items were endorsed by even 50% of any of the control groups (without NLUTD or UTI history), although those with a self-reported history of UTIs tended to endorse at higher rates while those without a history of UTIs endorsed at the lowest rates. These results suggest that the items are more descriptive for our target user group and less descriptive of the experiences of those

outside that group, particularly those without histories of urinary symptoms and UTIs. It is important to note that information about UTI history was self-reported, and we were unable to confirm category membership. We were also forced to combine the 1–5 and >5 UTIs voider groups into one “history of UTI” group in order to reach our target of 50 in this group, while for the IDC instrument we were able to identify individuals for the divergent validity groups in sufficient numbers who reported having lifetime histories of 1–5 UTIs and >5 UTIs. The convergent and divergent validity results tend to support our intention that the USQNB instruments reflect urinary (and other) symptoms that are associated with UTI.

As with our previous report of intermittent catheter users [21], the national samples queried participants about each item within the previous year (e.g., “Did you experience, in the past year,

**Table 4.** USQNB-V endorsement rates by group (Voiders with NLUTD, and convergent/ divergent groups).

Symptom	NLUTD/ SCI Void	NLUTD/ MS Void	NLUTD/MS or SCI Void, No UTIs	NLUTD MS/SCI No Void, + UTIs	No NLUTD, No UTIs	No NLUTD, + UTIs
n	104	405	118	137	50	55
A1: Bladder spasms	0.500	0.381	0.237	0.555	0.020	0.000
A2: Difficulty ambulating (walking)	0.480	0.721	0.642	0.750	0.340	0.345
A3: Worsening of motor function	0.406	0.468	0.432	0.331	0.180	0.200
A4: Increase in spasticity (not bladder spasms)	0.455	0.468	0.398	0.390	0.040	0.091
A5: Sweating	0.574	0.564	0.458	0.456	0.400	0.255
A6: Fever	0.426	0.371	0.229	0.426	0.160	0.164
A7: Difficulty thinking, mental foginess, forgetfulness	0.520	0.723	0.653	0.537	0.200	0.091
A8a: Autonomic Dysreflexia <b>**SCI ONLY**</b>	0.434		0.233	0.472	0.000	0.000
A8b: Multiple Sclerosis Flare-up (relapse, attack) <b>**MS ONLY**</b>		0.658	0.571	0.552	0.000	0.000
B1.1: Blood in urine	0.343	0.195	0.042	0.401	0.140	0.073
B1.2: Increased frequency of urination	0.686	0.899	0.814	0.628	0.400	0.182
B1.3: Urinary urgency	0.559	0.723	0.695	0.394	0.480	0.218
B1.4: Urge incontinence	0.725	0.802	0.593	0.577	0.240	0.182
B1.5: Sensation of urgency/empty bladder	0.353	0.388	0.347	0.234	0.160	0.109
B1.6: Inability to fully empty bladder	0.618	0.624	0.432	0.285	0.040	0.091
B1.7: Small release of urine	0.696	0.827	0.754	0.299	0.440	0.273
B1.8: Weak urine stream	0.500	0.582	0.432	0.234	0.060	0.127
B2.1: Bad, foul or stronger smelling urine	0.716	0.588	0.415	0.781	0.280	0.182
B2.2: Dark urine	0.718	0.528	0.373	0.664	0.400	0.200
B2.3: Cloudiness	0.721	0.588	0.263	0.839	0.300	0.073
B2.4: Sediment in discharge	0.412	0.151	0.034	0.518	0.020	0.018
C1: Malaise, feeling unwell	0.745	0.681	0.508	0.650	0.380	0.236
C2: Fatigue	0.833	0.950	0.907	0.891	0.600	0.418
C3: Chills	0.594	0.443	0.314	0.574	0.160	0.200
C4: Nausea	0.480	0.431	0.297	0.368	0.400	0.164
C5: Vomiting	0.270	0.213	0.144	0.213	0.300	0.145
C6: Muscle weakness	0.598	0.814	0.729	0.606	0.280	0.273

NOTES: Values in the table are rates, so 0.50 = 50.0% of the sample endorsed that item. USQNB-V: Urinary Symptom Questionnaire for Neurogenic Bladder or Neurogenic Lower Urinary Tract Dysfunction – for voiding; SCI = spinal cord injury; MS = multiple sclerosis; NLUTD = neurogenic lower urinary tract dysfunction or neurogenic bladder. UTI = urinary tract infection. “+UTI” = self report of at least one UTI diagnosis; “no UTI” = self report of no UTI diagnoses in lifetime. A = Clinically Actionable symptoms; B1 = Bladder Specific symptoms; B2 = Urine Quality symptoms; C = Constitutional symptoms.

increased cloudiness of urine?”), even though the intended use of the instrument is for a much shorter time window, such as one or two weeks. While we have observed different endorsement rates by week over an 18-month period for the USQNB-IC items [18], and over a 12-month period when assessed biweekly for the USQNB-V [26] and USQNB-IDC [27], the analysis and interpretability of these fluctuations is not yet clear. We have reported [18, 28] different methods of summarization (i.e., scoring) of the USQNB instruments that can be used together with our reports of the validity and reliability of all three instruments for continuing implementation of these surveys and their use by patients, clinicians, and investigators.

Although our national samples completing the USQNB-IDC and -V, like with the USQNB-IC [19], support conclusions of validity and

reliability [20], these national surveys attracted respondents who are quite heterogeneous. That is, respondents were *not excluded* if they had any of the following conditions: (1) known genitourinary pathology beyond neurogenic bladder (i.e., vesicoureteral reflux, bladder or kidney stones, etc.); (2) use of prophylactic antibiotics; (3) instillation of intravesicular agents to reduce UTI (i.e., gentamycin); (4) psychologic or psychiatric conditions influencing the ability to follow instructions; and (5) participation in another study in which results would be confounded. Thus, all three instruments derive validity and reliability evidence from national samples that are possibly more similar to a typical clinical sample than to a typical clinical trial/efficacy study sample. Importantly, these characteristics were all exclusion criteria for the focus groups from whom the items on all instruments were originally obtained;



since it is possible that our national sample included respondents with some of these potentially confounding factors, it tends to strengthen the argument from this evidence of the face, content, and convergent validity of the instruments. However, everyone who agreed to participate was self-selected (just as our focus groups were in developing the instruments in the first place). We did not inquire about facilitated voiding techniques (Valsalve, Crede, etc), as there is no standardized way to interpret patient report of these techniques. Similarly, we did not specify a type of indwelling catheter and responses on the USQNB-IDC from the target group would have included individuals with both indwelling urethral and suprapubic catheters. While research applications of these instruments would have possibly more homogeneous respondents than our samples were, the validity evidence comes from heterogeneous users that are more consistent with clinical practice.

With the validity (this paper) and reliability [20] evidence supporting the further use and implementation of these instruments in clinical, research, and self-management contexts, greater depth of study into the existence of bias is possible, including recall bias. This study used a 12-month recall period to ensure we

could capture real face validity evidence from our national target sample. However, the instruments are intended for use over a much shorter time frame (1–2 weeks), which will limit recall bias.

Our approach to validity evidence leverages formal psychometric criteria [23] and consensus-based approaches [21] to the documentation of validity in clinical assessment. Both include considerations of associations with gold standards (COSMIN criterion validity) or decisions (Geisinger’s “consequences of testing”), and because the purpose of the USQNB instruments is to strengthen the representation of the patient’s experience of urinary signs and symptoms in the diagnosis of UTI-and possibly, sub-clinical treatable levels of symptoms- and improve our understanding of NLUTD, neither of these aspects of validity could be included in this study. We used the same methods [13] to develop the two USQNB instruments described here as were used for the first USQNB, for intermittent catheter users (USQNB-IC [19]). For these three independent studies, similar sources of validity evidence were obtained, even though the instruments have different items. Our team continues to analyze these rich data sets in hopes of continuing to move research and clinical care forward.

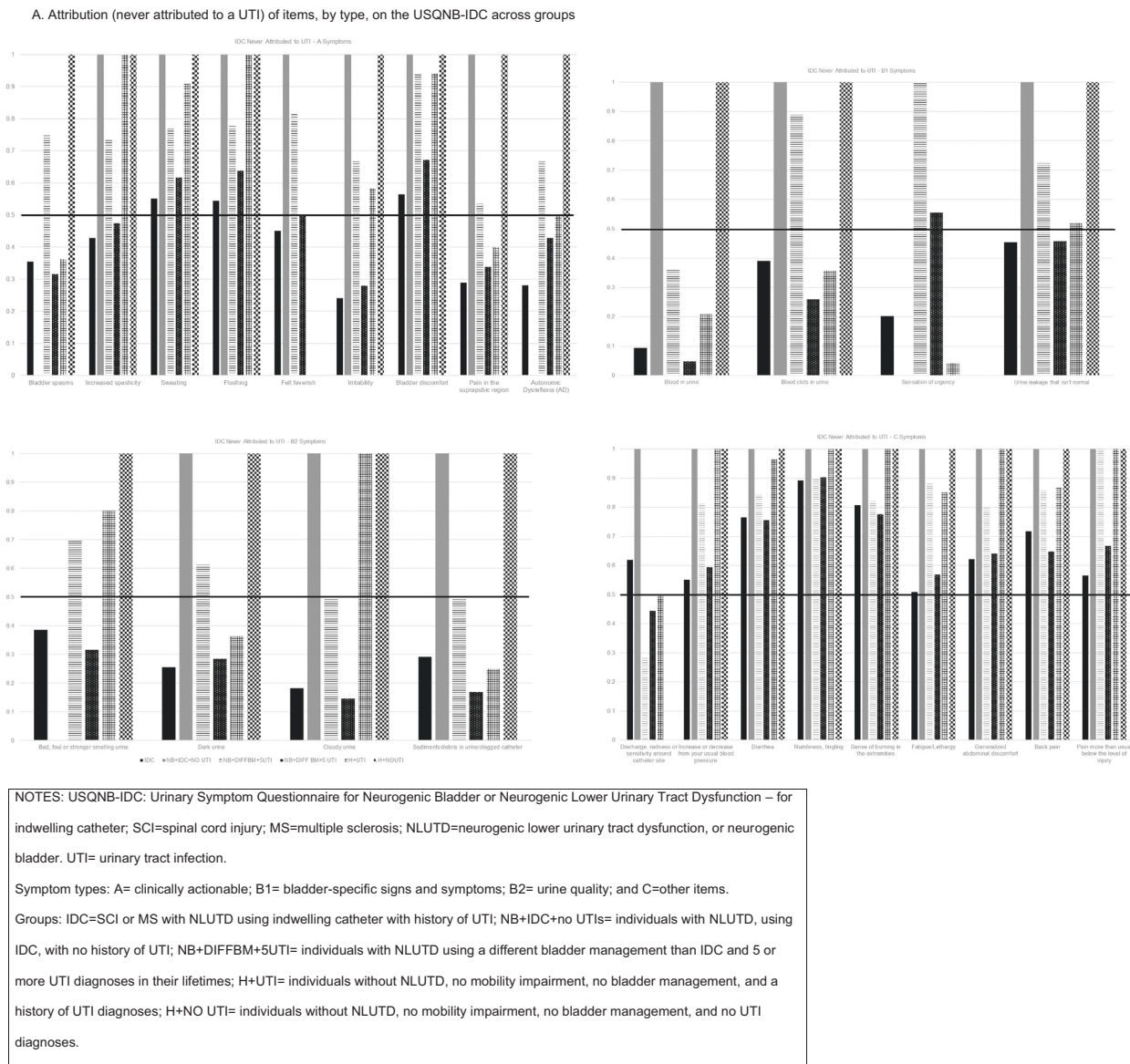
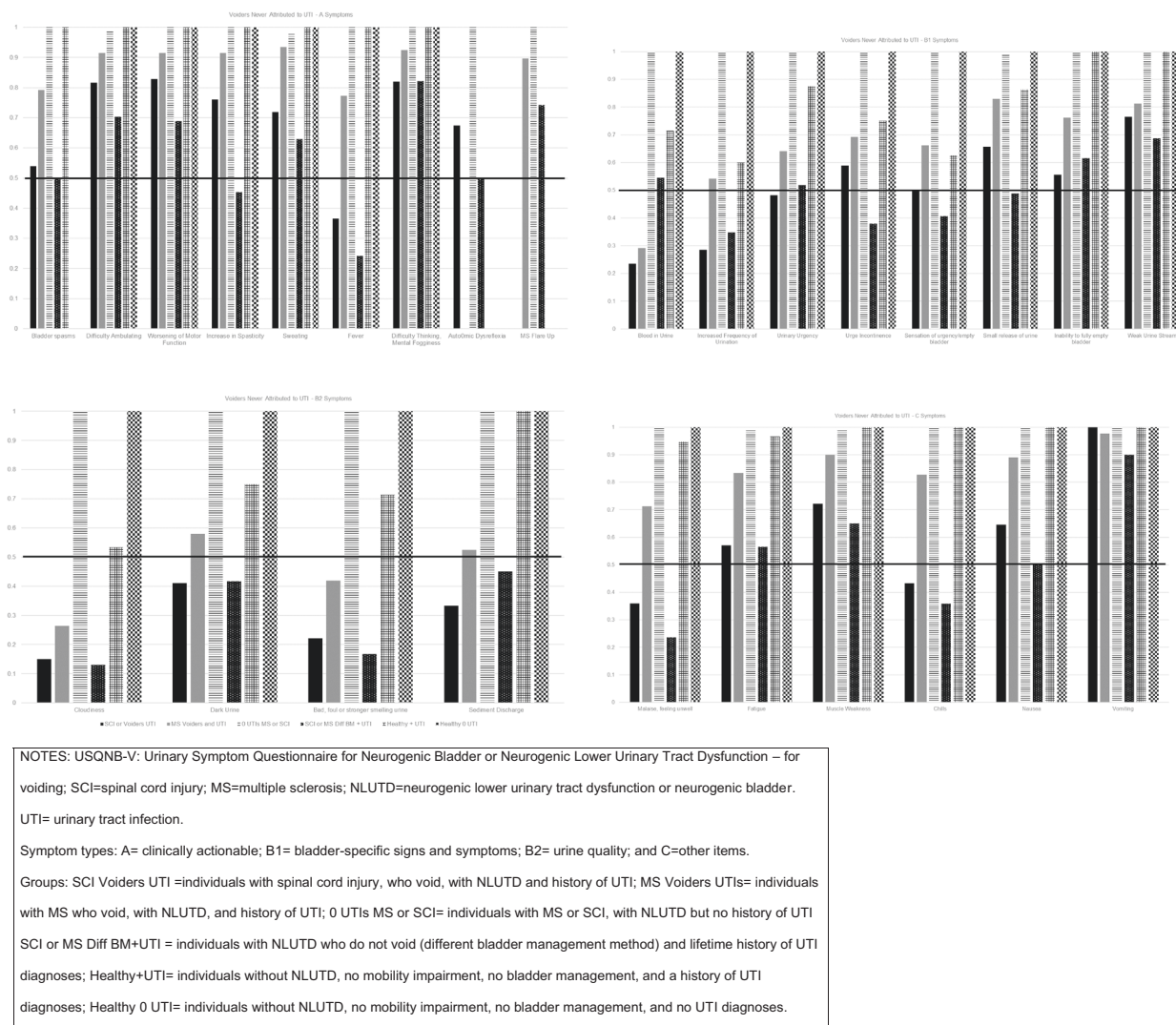


Fig. 2 (Continued).

B. Attribution (never attributed to a UTI) of items, by type, on the USQNB-V across groups



**Fig. 2** Percentage of those endorsing each item who never attributed it to a UTI. **A** Attribution (never attributed to a UTI) of items on the USQNB-IDC across groups. **B** Attribution (never attributed to a UTI) of items on the USQNB-V across groups.

## CONCLUSIONS

The *patient-centered* patient reported outcomes discussed here represent our efforts at “valuing the patient perspective” and maintaining a “culture of patient centeredness in research”. It is important to keep in mind that formal measurement properties that are of greatest interest [21, 24] in health outcomes (reliability and validity) are defined based on the uses to which the instrument will be put [22, 23, 25, 29]. Our results suggest that, like we reported for the USQNB-IC [19], the USQNB-V and USQNB-IDC provide a valid, coherent, and comprehensive view of the patient’s experience of urinary symptoms along the continuum of NLUTD. The three instruments were developed independently but have yielded the same level and types of validity evidence which strengthens our confidence in the results. A new paper outlining all three USQNB instruments, together with COSMIN-appropriate scoring information, is in preparation. These instruments are clinically relevant because patient-reported urinary symptoms are bothersome and common, but not the same as the symptoms identified by authoritative guidelines for diagnosis of UTI [13]. Having an instrument that allows for measurement of both patient- and clinician-determined symptoms that are potentially

related to UTI for individuals with NLUTD specifically, and depending on bladder management, we can begin to actually differentiate those symptoms which are definitely, probably, or unlikely related to UTI among those with NLUTD. Our team continues to work to improve treatment and research into UTI and bothersome urinary symptoms (e.g., Tractenberg et al. 2020 [18]), as well as to help promote antibiotic stewardship by focusing treatments where antibiotics would be most likely to be effective.

## DATA AVAILABILITY

Data supporting the results reported in the article can be found in Files section of the Open Science Framework Project page, <https://osf.io/9e7yp/>

## REFERENCES

- Kavanagh A, Baverstock R, Campeau L, Carlson K, Cox A, Hickling D, et al. Canadian Urological Association guideline: Diagnosis, management, and surveillance of neurogenic lower urinary tract dysfunction - Full text. CUAJ [Internet]. 2019 Feb [cited 2020 May 12];13. Available from: <https://cuaj.ca/index.php/journal/article/view/5912>
- Dorsner PT, McIntosh PM. Neurogenic bladder. Adv Urol. 2012;2012:1–16.

3. Benevento BT, Sipski ML. Neurogenic bladder, neurogenic bowel, and sexual dysfunction in people with spinal cord injury. *Phys Ther.* 2002;82(Jun):601–12.
4. Cardenas DD, Hoffman JM, Kirshblum S, McKinley W. Etiology and incidence of rehospitalization after traumatic spinal cord injury: a multicenter analysis. *Arch Phys Med Rehabil.* 2004;85(Nov):1757–63.
5. Liu C-W, Attar KH, Gall A, Shah J, Craggs M. The relationship between bladder management and health-related quality of life in patients with spinal cord injury in the UK. *Spinal Cord.* 2010 ;48(Apr):319–24.
6. Stillman MD, Barber J, Burns S, Williams S, Hoffman JM. Complications of spinal cord injury over the first year after discharge from inpatient rehabilitation. *Arch Phys Med Rehabilitation.* 2017;98(Sep):1800–5.
7. Savic G, DeVivo MJ, Frankel HL, Jamous MA, Soni BM, Charlifue S. Causes of death after traumatic spinal cord injury—a 70-year British study. *Spinal Cord.* 2017;55 (Oct):891–7.
8. Hooton TM, Bradley SF, Cardenas DD, Colgan R, Geerlings SE, Rice JC, et al. Diagnosis, prevention, and treatment of catheter-associated urinary tract infection in adults: 2009 International Clinical Practice Guidelines from the Infectious Diseases Society of America. *Clin Infect Dis: Off Publ Infect Dis Soc Am.* 2010;50 (Mar):625–63.
9. Madden-Fuentes RJ, McNamara ER, Lloyd JC, Wiener JS, Routh JC, Seed PC, et al. Variation in definitions of urinary tract infections in spina bifida patients: a systematic review. *Pediatrics* 2013 ;132(Jul):132–9.
10. Everaert K, Lumen N, Kerckhaert W, Willaert P, van Driel M. Urinary tract infections in spinal cord injury: prevention and treatment guidelines. *Acta Clin Belg.* 2009;64 (Aug):335–40.
11. Naber KG, Bergman B, Bishop MC, Bjerklund-Johansen TE, Botto H, Lobel B, et al. EAU guidelines for the management of urinary and male genital tract infections. Urinary Tract Infection (UTI) Working Group of the Health Care Office (HCO) of the European Association of Urology (EAU). *Eur Urol.* 2001;40(Nov):576–88.
12. Tractenberg RE, Groah SL, Rounds AK, Davis EF, Ljungberg IH, Frost JK, et al. Clinical Profiles and Symptom Burden Estimates to Support Decision-Making Using the Urinary Symptom Questionnaire for People with Neurogenic Bladder (USQNB) using Intermittent Catheters. *PM R.* 2021;13:229–40. <https://doi.org/10.1002/pmrj.12479>.
13. Tractenberg RE, Garver A, Ljungberg IH, Schladen MM, Groah SL. Maintaining primacy of the patient perspective in the development of patient-centered patient reported outcomes. *PLOS ONE.* 2017 ;12(Mar):e0171114.
14. Tate DG, Wheeler T, Lane GI, Forchheimer M, Anderson KD, Biering-Sorensen F, et al. Recommendations for evaluation of neurogenic bladder and bowel dysfunction after spinal cord injury and/or disease. *J Spinal Cord Med.* 2020;43(Mar):141–64.
15. Patel DP, Elliott SP, Stoffel JT, Brant WO, Hotaling JM, Myers JB. Patient reported outcomes measures in neurogenic bladder and bowel: A systematic review of the current literature: Patient Reported Outcomes in Neurogenic Bladder and Bowel. *NeuroUrol Urodyn.* 2016;35(Jan):8–14.
16. Best KL, Ethans K, Craven BC, Noreau L, Hitzig SL. Identifying and classifying quality of life tools for neurogenic bladder function after spinal cord injury: A systematic review. *J Spinal Cord Med.* 2017;40(Sep):505–29.
17. Groah SL, Rounds AK, Ljungberg IH, Sprague BM, Frost JK, Tractenberg RE. Intravesical *Lactobacillus rhamnosus* GG is safe and well tolerated in adults and children with neurogenic lower urinary tract dysfunction: first-in-human trial. *Ther Adv Urol.* 2019;11:1756287219875594. <https://doi.org/10.1177/1756287219875594>.
18. Tractenberg RE, Groah SL, Frost JK, Rounds AK, Davis E, Ljungberg IH, et al. Effects of Intravesical *Lactobacillus Rhamnosus* GG on Urinary Symptom Burden in People with Neurogenic Lower Urinary Tract Dysfunction. *PM R.* 2021;13:695–706. <https://doi.org/10.1002/pmrj.12470>.
19. Tractenberg RE, Groah SL, Rounds AK, Ljungberg IH, Schladen MM. Preliminary validation of the Urinary Symptom Questionnaire for individuals with neuro-pathic bladder using intermittent catheterization (USQNB-IC): A patient-centered patient reported outcome. *PLOS ONE.* 2018;13(Jul):e0197568 <https://doi.org/10.1371/journal.pone.0197568>. eCollection 2018
20. Tractenberg RE, Frost JK, Rounds AK, Ljungberg IH, Schladen MM, Groah SL. Reliability of the Urinary Symptom Questionnaires for people with neurogenic bladder (USQNB) who void or use indwelling catheters. *Spinal Cord.* (In press).
21. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63(Jul):737–45.
22. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: American Educational Research Association; 2014. p.230.
23. Geisinger KF. Intended and unintended meanings of validity: some clarifying comments. *Assessment in Education: Principles. Policy Pract.* 2016;23(Apr):287–9.
24. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res.* 2018 ;27(May):1147–57.
25. Banta TW, Palomba CA. Assessment essentials: planning, implementing, and improving assessment in higher education. Second Edition. San Francisco: Jossey-Bass & Pfeiffer Imprints; 2015. p. 339.
26. Groah SL, Tractenberg RE, Rounds AK, Frost JK, Ljungberg IH. Urinary Symptoms and Urine Dipstick Self-Assessment are unassociated in individuals with SCI, MS, and NLUTD who Void. *Topics in Spinal Cord Injury Research.* (In press).
27. Rounds AK, Tractenberg RE, Groah SL, Frost JK, Ljungberg IH, Navia H, Pham CT. Urinary Symptoms are Unrelated to Leukocyte Esterase and Nitrite Among Indwelling Catheter Users (In-Revisions).
28. Tractenberg RE, Groah SL, Frost JK, Yumoto F, Rounds AK, Ljungberg IH. (In preparation). The spectrum of urinary symptoms among people with neurogenic lower urinary tract dysfunction (NLUTD). (In Preparation).
29. Lane S, Raymond MR, Haladyna TM, editors. Handbook of test development. Second edition. New York: Routledge; 2015.

## ACKNOWLEDGEMENTS

We gratefully acknowledge the contributions of Manon M. Schladen (Data curation, Investigation, and Acquisition of funding). Two funders supported this work: USQNB-V: Supported by NIDILRR grant #90IF0121 USQNB-IDC: Supported by Craig H. Neilsen Foundation grant # 385077

## AUTHOR CONTRIBUTIONS

RET: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing—review & editing. JFK: Formal analysis, Validation, Visualization, Writing—original draft, Writing—review & editing. FY: Formal analysis, Methodology, Validation, Visualization, Writing—original draft, Writing—review & editing. AKR: Data curation, Resources, Writing—review & editing. IHL: Project administration, Resources, and Acquisition of funding, Writing—review & editing. SLG: Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing—review & editing

## STATEMENT OF ETHICS

We certify that all applicable institutional and governmental regulations concerning the ethical use of human volunteers were followed during the course of this research.

## COMPETING INTERESTS

The authors declare no real or apparent conflicts of interest.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to R.E.T.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.