



Patient reported outcome measures and quality of life assessment in genital gender confirming surgery

Geolani W. Dy^{1#}, Ian T. Nolan^{2#}, James Hotaling³, Jeremy B. Myers³

¹Department of Urology, ²School of Medicine, New York University, New York, NY, USA; ³Department of Urology, University of Utah, Salt Lake City, UT, USA

Contributions: (I) Conception and design: All authors; (II) Administrative support: All authors; (III) Provision of study material or patients: All authors; (IV) Collection and assembly of data: IT Nolan, GW Dy; (V) Data analysis and interpretation: IT Nolan, GW Dy; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Geolani W. Dy. NYU Urology, 222 East 41st St., 11th Floor, New York, NY 10017, USA. Email: Geolani.dy@nyulangone.org.

Abstract: Transgender and gender nonbinary (TGNB) individuals may seek genital gender confirming surgery (GCS) as part of their transition. Outcomes of genital GCS may include gender congruence, sexual functioning and satisfaction, urinary symptoms, aesthetic satisfaction, and overall quality of life, among others. Despite a wide number of studies on results of vaginoplasty, metoidioplasty, phalloplasty, and other genital GCS, data regarding patient reported outcomes are limited. To date, there is no patient reported outcome measure (PROM) validated within the TGNB population to assess subjective outcomes of GCS. In this review, we aim to describe existing tools being utilized to report outcomes following GCS, assessing the merits and limitations of each.

Keywords: Transgender; body image; sex reassignment; patient-reported outcome; questionnaire; feminization; masculinization; quality of life (QoL); satisfaction

Submitted Nov 16, 2018. Accepted for publication Apr 23, 2019.

doi: 10.21037/tau.2019.05.04

View this article at: <http://dx.doi.org/10.21037/tau.2019.05.04>

Introduction

For transgender and gender nonbinary (TGNB) individuals, genital gender confirming surgery (GCS) (also known as gender affirming surgery or “bottom surgery”), aims to better align TGNB individuals’ genitourinary anatomy with their identified gender. Feminizing genital procedures may involve vaginoplasty, penectomy, clitoroplasty, vulvoplasty and orchiectomy. Masculinizing procedures include metoidioplasty, phalloplasty, scrotoplasty, colpectomy, testicular and penile prosthesis placement, hysterectomy and oophorectomy.

Subjective outcomes such as degree of dysphoria, overall quality of life (QoL), and individuals’ satisfaction are among the most important outcomes of any gender confirming operation, as these procedures are meant to allow TGNB individuals to live more comfortably in their bodies. Recent

systematic reviews on surgical techniques and outcomes of genital GCS reveal high overall effectiveness (1,2). However, these reviews are limited by the literature’s poor description of TGNB individuals’ subjective outcomes and satisfaction, which stems largely from the use of patient reported outcome measures that are not validated specifically for TGNB patients after GCS. In a systematic review and meta-analysis of vaginoplasty in transgender women, outcomes of overall satisfaction (93%), and satisfaction with functional outcomes (87%), aesthetic outcomes (90%), and sexual outcomes such as the ability to orgasm (70%) are described (3). Following phalloplasty and metoidioplasty in transgender men, urinary outcomes such as standing micturition (90% in metoidioplasty, 75% in phalloplasty), sexual outcomes including erogenous sensation (100% in metoidioplasty and 69% in phalloplasty) and penetration (51% in metoidioplasty, 43% in phalloplasty), and aesthetic

satisfaction (87% in metoidioplasty, 70% in phalloplasty) have been reported (4).

Data regarding patient-reported outcomes is generally limited, both by the size and low number of studies, and by the means by which data is collected. Patient-Reported Outcome Measures (PROMs) are questionnaires used to assess broader patient perceptions, experiences, and QoL and are ideally distributed to patients before and/or after an intervention. In GCS these may assess general satisfaction, overall QoL, diagnostic severity (e.g., degree of gender dysphoria or other comorbid mental health conditions like depression and anxiety), and specific aesthetic and specific functional (e.g., urinary or sexual) outcomes (5). Use of PROMs shifts attention toward patients' subjective experience with care and allows for more patient-centered evaluation of treatment efficacy. Lack of PROMs validated for use in TGNB individuals after GCS has limited our understanding of efficacy of individual treatments, our ability to compare relative efficacy of treatments within a group (i.e., phalloplasty versus metoidioplasty), our understanding of what TGNB individuals value regarding GCS pre- and post-operatively, therefore limiting our ability to improve techniques to meet their preferences (4,6).

Creating a *de novo* PROM using approved processes is labor intensive but would ultimately yield a questionnaire that organically addresses the needs of the population. The alternative, which has been adopted by all existing GCS studies that report subjective outcomes, is to use existing measures (either in whole or in part) that are not validated specifically for transgender people undergoing GCS. For example, use of measures validated for TGNB people but not after GCS are likely to be confounded by aspects of QoL that are unlikely to be directly affected following surgery, such as family acceptance or interpersonal relationships; and use of measures that are validated after breast surgery but not for TGNB people are likely to overlook unique components of QoL, such as being identified as the correct gender or relieving body dysmorphia.

To date, PROMs used in GCS studies include: tools validated in TGNB individuals but not specific for the effects of GCS, tools validated to report functional outcomes (urinary, sexual, and aesthetic) but not in TGNB individuals, tools with incomplete validation for use in TGNB individuals after GCS, and *ad-hoc* tools without validation for any purpose (2). Early efforts to create validated PROMs for GCS are underway (7). This review will discuss specific PROMs used to measure subjective

outcomes of GCS, considering the merits and limitations of each. *Table 1* summarizes the various assessment tools described in this review.

Proms used to assess psychosocial outcomes

General QoL measures

Several general QoL PROMs intended for use in the general population and not specifically in TGNB people, have been used to measure QoL changes following GCS (2). These include the World Health Organization (WHO) QoL-100, 36-Item Short Form Health Survey (SF-36), Symptom Checklist 90 (SCL-90), and Satisfaction With Life Scale (SWLS), among others (8,26-28). These instruments may provide some insight into QoL before and after surgery, but since they are not condition specific other life-events, such as social or medical changes, can negatively or positively affect individuals' responses and the outcomes of the survey may have little to do with GCS (26). General QoL measures do not adequately assess specific psychosocial or medical factors that affect QoL in transgender people specifically, such as gender dysphoria; higher rates of anxiety, depression, and substance use; higher rates of poverty and unemployment; and higher rates of HIV (29,30). Therefore, it would be ideal for QoL measurements in GCS to use tools specifically designed to capture the experience of TGNB individuals undergoing GCS.

Transgender/non-binary specific measures

Several PROMs have been validated in transgender populations to measure psychosocial outcomes like QoL, gender dysphoria, and gender congruence. These include the Transgender Congruence Scale (TCS), the Biographical Questionnaire for Transvestites and Transsexuals (BQTT), the Utrecht Gender Dysphoria Scale (UGDS) and its non-binary adaptation, the UGDS-Gender Spectrum (UGDS-GS), the Gender Identity/Gender Dysphoria Questionnaire for Adults and Adolescents (GIDYQ-AA), and the Body Image Scale (BI-1) (9-11,13,31). The TCS, UGDS-GS, or GIDYQ-AA are likely the most useful in modern GCS research, although all existing measures have notable limitations.

TCS

The TCS is a 12-item tool in which subjects rate various

Table 1 Existing PROMs used in measuring TGNB individuals' quality of life/psychiatric outcomes, urinary outcomes, sexual outcomes, and aesthetic outcomes, as they relate to gender confirming surgery

Tool	Validated populations	QoL/psychiatric outcomes	Urologic outcomes	Sexual outcomes	Aesthetic outcomes	Description	Strengths	Weaknesses
Satisfaction with Life Scale (SWLS) [Diener, 1985 (8)]	All patients generally	x				5-item questionnaire with 7-point Likert scale assessing general satisfaction with life	Validated to measure general satisfaction with life	An enormous number of factors are likely to influence the result of this measure
Body Image Scale (BI-1) [Lindgren, 1975 (9)]	TG (excluding GNB)					30-item questionnaire with 5-point Likert scale rating satisfaction with anatomic parts, and question regarding desire to change each part medically or surgically. Higher score indicates greater dissatisfaction. Three of 30 domains could be affected by bottom surgery (scrotum/vagina, penis/clitoris, testicles/ovaries/uterus)	Validated in transgender patients. Can be used pre- and post-operatively to assess overall satisfaction with individual body parts	Not validated in GNB populations Not specific to gender-related body dysphoria Changes in any one body part are unlikely to meaningfully change a patient's score At least 10/30 (breasts, arms, muscles, facial hair, biceps, hair, voice, body hair, chest, appearance) are likely to be affected by hormone therapy or voice therapy
Biological Questionnaire for Transvestites and Transsexuals (BQTT) [Verschoor, 1988 (10)]	Not validated	x	x			250-item questionnaire referring to sociodemographic information, gender development during adolescence and adulthood, preadolescent gender behavior, "transvestite practice", sexuality, and medical antecedents, stable sexual relationship, sex of partner, sexual satisfaction with the partner, frequency of orgasm, frequency of masturbation, and frequency of sexual arousal	Best used as a general QoL measure	Not validated Includes a great number of questions about psychosocial factors not likely to be meaningfully affected by surgery Very long, cumbersome to dispense
Utrecht Gender Dysphoria Scale (UGDS) [Cohen-Kettenis, 1997 (11)]	TG (excluding GNB)	x		x		12-item scale (separate scales for M and F) that assesses gender dysphoria. Examples, "I feel a continuous desire to be treated as a man/woman" and "Every time someone treats me as a woman/man I feel hurt."	Five of 12 items relate to GCS directly (UGDS-M): "I feel unhappy because I have a male body", "The idea that I will always be a boy/man gives me a sinking feeling", "I hate myself because I am a boy/man", "I dislike urinating in a standing position", "I dislike having erections."	Necessarily depends on a gender binary, excludes GNB patients Not specifically validated for effects of GCS Likely to be affected by non-surgical treatments that also are effective at reducing gender dysphoria Includes questions not likely to change as a result of any kind of gender affirmation treatment

Table 1 (continued)

Table 1 (continued)

Tool	Validated populations	QoL/psychiatric outcomes	Urologic outcomes	Sexual outcomes	Aesthetic outcomes	Description	Strengths	Weaknesses
Utrecht Gender Dysphoria Scale-Gender Spectrum (UGDS-GS) [McGuire, 2016 (12)]	Adapted from UGDS for gender spectrum	x			x	18-item scale that assesses gender dysphoria. Example adapted question from UGDS, "Every time someone treats me like my assigned sex I feel hurt."	Has been adapted but not validated for GNB patients	Limitations mirror those of UGDS
Gender Dysphoria Questionnaire for Adults and Adolescents (GIDYQ-AA) [Degracias, 2007 (13)]	TG (excluding GNB)	x				27-item questionnaire with a 5-point Likert scale, including 13 subjective, 9 social, 3 somatic, 2 sociologic items. Measures gender dysphoria ("I feel more like the opposite sex", "I don't feel like I am of my current sex", "Strangers treat me like a member of the opposite sex"), as well as those that are likely to be affected by GCS ("Wise for sex-reassignment surgery", "Felt more like the opposite sex", "Thought of self as opposite sex", "Not feeling like they are of their current sex", "Satisfaction with current sex", "Dreaming of being the opposite sex", "Upset reuse of current sex on official forms", "Anatomic dysphoria", "Strangers treat person as opposite sex", "Presented self as opposite sex at parties", "Thought of sex as current sex", "Friends or relatives treat person as opposite sex", "Presented self as opposite sex at work or school", "Having to work at being their current sex", "Comfort with using restrooms of biological sex")	Many [16] items are likely to be affected by gender affirming surgery: "Wise for sex-reassignment surgery", "Felt more like the opposite sex", "Thought of self as opposite sex", "Not feeling like they are of their current sex", "Satisfaction with current sex", "Dreaming of being the opposite sex", "Upset reuse of current sex on official forms", "Anatomic dysphoria", "Strangers treat person as opposite sex", "Presented self as opposite sex at parties", "Thought of sex as current sex", "Friends or relatives treat person as opposite sex", "Presented self as opposite sex at work or school", "Having to work at being their current sex", "Comfort with using restrooms of biological sex"	Not validated in GNB populations Several items are better indicators of transgender identity rather than GD. Only a few questions [5] ask about feelings regarding transgender identity and GD Some questions [7] appear to conflate "sex" with "gender"
King's Health Questionnaire (KHQ) [Kelleher, 1997 (14)]	Cisgender women	x	x			21-item questionnaire with a 4-point Likert scale. Measures the psychosocial effects of urinary incontinence: psychical, social, and personal limitations; emotional problems; and sleep or energy disturbance	Validated measure likely to capture several aspects of urologic health post-vaginoplasty	Not validated in TG/NB populations

Table 1 (continued)

Table 1 (continued)

Tool	Validated populations	QoL/psychiatric outcomes	Urologic outcomes	Sexual outcomes	Aesthetic outcomes	Description	Strengths	Weaknesses
Sheffield Pelvic Organ Prolapse Quality of Life Questionnaire (SPS-Q) [Bradshaw, 2006 (15)]	Cisgender women		x	x	x	25-item questionnaire with a 4-point Likert scale. Measures presence and severity of urinary, bowel, and sexual function symptoms associated with vaginal prolapse	Validated measure likely to capture several aspects of urologic and sexual health post-vaginoplasty	Not validated in TGNB populations
Body Image Scale for Women (BISFW) [Taylor, 1994 (16)]	Cisgender women			x		22-item scale measuring interest, activity, and satisfaction. Meant to be comparable in form and content to the BSFQ for Men (17)	Includes sexual activity, subjective factors like satisfaction with sex life and interpersonal consequences of sexual health, and details of sexual dysfunction like frequency of dyspareunia, lack of lubrication, anorgasmia, vaginal tightness, incontinence, and vaginal infection	Not validated in TGNB populations Will not optimally capture psychosocial features of sexual health specific to TGNB populations
Female Sexual Function Index (FSFI) [Rosen, 2000 (18)]	Cisgender women			x		19-item scale measuring desire, arousal, lubrication, orgasm, satisfaction, and sexual pain or discomfort	More focus on individual components of sexual health	Not validated in TGNB populations Less focus on interpersonal results of sexual health
Sexual Functioning Index-Gender Spectrum (SF-GS) [Spencer, 2017 (19)]	-			x		19-item questionnaire assessing 5 domains: sexual activity, alone or with a partner; sexual interest; physical arousal; orgasm; and sexual pain or discomfort	Adapted for gender spectrum. Most questions are applicable for pre- and post-operative comparisons for any form of GCS	Not validated
Female Genital Self Image Score (FGSIS) [Hebernick, 2010 (20)]	Cisgender women			x		7-item scale measuring satisfaction with genitals globally, and specifically regarding sexual function and aesthetics	Best used to measure overall satisfaction with genitals	Not validated in TGNB populations Less focus on details of what contributes to sexual function or aesthetic results
Sexual Desire Inventory (SDI) [Spector, 1996 (21)]	Cisgender women			x		14-item scale measuring sexual desire	Excellent for measuring sexual desire	Not validated in TGNB populations Poor measurement of parameters other than sexual desire

Table 1 (continued)

Table 1 (continued)

Tool	Validated populations	QoL/psychiatric outcomes	Urologic outcomes	Sexual outcomes	Aesthetic outcomes	Description	Strengths	Weaknesses
Amsterdam Hyperactive Pelvic Floor Scale for Women (AHPFS-W) [Postma, 2013 (22)]	Cisgender women		x	x		30-item scale measuring stress symptoms [6], lower urinary tract symptoms [4], irritable bowel symptoms [4], rectal symptoms [3], and urinary tract infection symptoms [2]. Originally validated in women following sexual trauma	May capture several sexual and urinary health concerns important to transgender women	Not validated in TGNB populations Different psychosocial contributors to sexual dysfunction in transgender women than in trauma
Morrison, 2015 (23)	Not validated		x	x	x	<i>Ad-hoc</i> assessment of mucorrhea, dyspareunia, increased need for dilation, need for lubrication, bad odor, ability to achieve orgasm, and satisfaction with appearance, sexual function, postoperative recover, and overall satisfaction	Several outcomes highly specific to transgender women after intestinal vaginoplasty, such as need for dilation and mucorrhea	Not validated
Wierckx, 2011 (24)	Not validated			x		<i>Ad-hoc</i> questionnaire compares sexual function during masturbation and sexual activity before and after vaginoplasty: frequency of masturbation, ability and frequency of orgasm, quality of orgasm, use of vagina and clitoris during masturbation, sexual arousal	Several outcomes highly relevant to transgender women after vaginoplasty, such as use of neoclitoris and neovagina and ability to orgasm through sex or masturbation	Not validated
IPSS [Barry, 2017 (25)]	Cisgender men		x			7-item scale with 6-point Likert, measuring the frequency of urinary symptoms commonly associated with benign prostate hyperplasia, including retention, frequency, urgency, weak stream, straining, and nocturia	Validated to measure obstructive, irritative lower urinary tract symptoms	Does not include specific urologic outcomes relevant to phalloplasty, metoidioplasty, colpectomy, including ability to stand while voiding, split/spraying stream, pelvic or perineal pain with voiding

UGDS, Utrecht Gender Dysphoria Scale; UGDS-M, Utrecht Gender Dysphoria Scale-Male version; QoL, quality of life; GCS, gender confirming surgery; GNB, gender nonbinary; TGNB, transgender and gender nonbinary.

aspects of their comfort with their gender identity and expression on a 5-point Likert scale (31). The TCS has two domains: appearance congruence (9 items) and gender identity acceptance (3 items). Of these, the nine questions assessing appearance congruence are very likely to be relevant to GCS. They assess both opinions (“My outward appearance represents my gender identity”, “I feel that my mind and body are consistent with one another”) and also emotional reactions (“I am generally comfortable with how others perceive my gender identity when they look at me”, and “I am happy with the way my appearance expresses my gender identity”) which affect gender dysphoria and are also heavily influenced by GCS.

Unlike other measures, which may be validated in transgender individuals, the TCS is also validated in non-binary and non-conforming identities. The majority of PROMs used in GCS research, if valid in transgender populations at all, are not specifically validated to encompass the diverse range of gender identities that can be classified broadly as “TGNB”, and thereby are likely to give suboptimal data for up to half of TGNB people (29,32).

The main limitation of TCS use in GCS is that it is not specifically validated to measure the effects of surgery. Many items in the appearance congruence domain are likely to be influenced by non-surgical gender affirming treatments like hormone therapy (33). As there are no questions assessing satisfaction or congruence regarding specific body parts, scores are also likely influenced by gender affirmation targeting body parts other than those that are targeted by the surgery being assessed.

Additionally, the three items in the gender identity acceptance factor are less likely to be affected by GCS. The three items: “I am not proud of my gender identity”, “I am happy that I have the gender identity that I do”, and “I have accepted my gender identity”, are likely to affect overall QoL and may reflect gender dysphoria, but are less likely to be affected by GCS. Inclusion of these items could mute the apparent effects of surgery on this scale.

UGDS, UGDS-GS

The UGDS is a 12-item scale validated amongst transgender men and women to assess the subjective feelings individuals have regarding their gender identity (11). University of Minnesota’s UGDS-GS is an 18-item adaptation of the UGDS to include the gender non-binary population (12). These statements include, “I prefer to behave like my affirmed gender”, and “Every time someone treats me like my assigned sex I feel hurt”. These scales would be most

useful in assessing overall gender dysphoria-related QoL in TGNB individuals.

Compared with the TCS, the UGDS-GS has a lower proportion (4 of 18) of items that are likely to be directly affected by GCS. While none of the QoL measures that are commonly used are validated in TGNB people after GCS, this tool is probably sensitive to the effects of GCS. Likely GCS-dependent questions include: “It is uncomfortable to be sexual in my assigned sex”, “The bodily functions of my assigned sex are distressing for me (i.e., erection, menstruation)”, “I feel hopeless if I have to stay in my assigned sex”, and “I feel unhappy because I have the physical characteristics of my assigned sex”.

There are several shortcomings to using the UGDS in GCS-specific research. Most importantly, it is not specifically validated for effects of genital GCS or GCS generally. While four questions do pertain directly to GCS, 14 do not, and even those questions that are most related to GCS are likely to be influenced by other social and medical life events that may also affirm an individual’s gender. The items which seem to evaluate subjects for transgender identity are unlikely to be affected by any kind of gender confirming intervention. For example, answers to questions about dissatisfaction with having to live in a gender role congruent with the sex assigned at birth theoretically should not change as a result of gender affirmation therapies.

GIDYQ-AA

The GIDYQ-AA may be useful if slightly more objective data is desired, and may provide more granular detail about exactly which factors influence changes in gender dysphoria as a result of treatment (13). It is a 27-item tool that assesses the degree of gender dysphoria as well as the salience or frequency of events that are likely to cause gender dysphoria. About half the questions [13] ask about subjective feelings related to gender dysphoria. The others ask about social [9], somatic [3], and sociolegal [2] items that are likely to be related to gender dysphoria. For example, questions like “Strangers treat me like a member of the opposite sex” and “Having to work at being my current sex” give insight into how an individual “passes” as their affirmed gender in in the public sphere, which could be influenced by GCS.

Like the UGDS, about half [16] of the items in the GIDYQ-AA are likely to be directly affected by GCS. Many items assess outcomes that are direct goals of many TGNB individuals who undergo GCS, including “Felt more like the opposite sex”, “Thought of self as opposite sex”, “Anatomic dysphoria”, “Strangers treat person as opposite

sex”, “Friends or relatives treat person as opposite sex”, and “Comfort with using restrooms of biological sex”.

However, the GIDYQ-AA has limitations if applied to GCS studies. Several of the items in this tool better assess TGNB identity rather than gender dysphoria or the effects of GCS. For example, items such as “I feel more like the opposite sex”, “I don’t feel like I am of my current sex”, and “I have undergone efforts to change my legal sex”, do not necessarily reflect any dysphoria. They are also unlikely to be changed by GCS.

The GIDYQ-AA also may be confusing to subjects due terminology around gender and sex. Some questions [7] conflate “sex” with “gender”: For example, “Unhappiness with current sex”, and “comfort with using restrooms of biological sex”, may imply that sex is changed by GCS, but it is not. If this measure should be used, its language should be updated (34). The term “current sex” should be changed to “gender” or “outward gender expression”, and “biologic sex” should be changed to “sex assigned at birth”, depending on the context of each question. However, doing so would require further validation.

BI-1

The BI-1 is unlikely to be a useful index of changes in gender dysphoria following GCS for a number of reasons (9). It is a 30-item questionnaire asking subjects to rate dissatisfaction with various parts of the body on a 5-point Likert scale. However, measuring bodily dissatisfaction is a poor proxy for gender dysphoria, as bodily dissatisfaction may be the result of any number of factors beyond gender. Use in GCS is limited because the tool assesses dissatisfaction with bodily features that are unlikely to be changed by surgery, including stature and weight. Additionally, the tool’s broad assessment of bodily dissatisfaction is unlikely to reflect changes from GCS, which typically target a single body part and would therefore only change the response to 1 out of 30 items. Furthermore, only 3 of the 30 domains could be feasibly affected by genital GCS: “scrotum/vagina”, “penis/clitoris”, and “testicles/ovaries/uterus”.

BQTT

The BQTT has been used widely, although it has no formal validation. Its utility in GCS is somewhat limited due to its length, lack of specificity for GCS, and exclusion of nonbinary people (10). At 250 questions in length, full administration of this tool would likely be cumbersome and subjects may abort the questionnaire before completing

it. Furthermore, the questionnaire addresses a number of factors not likely to be meaningfully changed by GCS. These include a wide array of topics including sociodemographic information, gender development during adolescence and adulthood, preadolescent gender behavior, “transvestite practice”, sexuality, medical antecedents, stability of sexual relationship, sex of partners, sexual satisfaction with partner, frequency of orgasm, frequency of masturbation, and frequency of sexual arousal. As with most PROMs used in GCS research, the BQTT is not intended for use in nonbinary individuals, which limits its utility for up to half of TGNB people. Moreover, the terms “transvestite” and “transsexual” are considered outdated terms which are not appropriate for describing the TGNB community.

PROMs used to assess specific functional outcomes

PROMs used to assess specific functional outcomes of GCS include measures partially validated to measure outcomes of GCS, as well as measures that are well-validated to measure their respective outcomes in the general population (2). Many PROMs validated for various functional outcomes, but not validated to measure these specific objectives and outcomes in GCS have been used, and are limited in that they likely do not encompass symptoms specific to the genitourinary surgery (e.g., urinary spraying, ability to have penetrative intercourse) or the gender-confirming objectives of the individuals who undergo GCS (e.g., desire for standing micturition or penetrative intercourse).

Those measures that were designed specifically for use in TGNB people are commonly *ad-hoc* instruments with no validation process, which will not be discussed due to their low likelihood of relevance for other studies (2). *Ad-hoc* instruments involve questions that were devised by study investigators to assess outcomes, but have not undergone a validation process. These questions can yield very important information, such as simply asking “are you satisfied with the surgery?” However, they are prone to problems, such as asking “would you do the surgery again?” where many individuals will answer “no” because they have already had the surgery. Tools that underwent an incomplete validation process will be discussed, with the understanding that they are psychometrically suboptimal, as they typically included 1 or fewer steps of the suggested validation process, which includes (I) item generation either from literature review, expert opinion, or patient interviews,

(II) evaluation of psychometric properties, followed by (III) item reduction to minimize the length of the questionnaire.

Urinary outcomes

Urinary outcomes in transgender females following vaginoplasty have been measured using function-specific measures not validated for TGNB individuals, namely the Sheffield Pelvic Organ Prolapse Quality of Life Questionnaire (SPS-Q) and the King's Health Questionnaire (KHQ) (2,14,15). Both may be of utility in post-vaginoplasty assessments, depending on the desired outcomes to be measured. The SPS-Q is a 25-item questionnaire with a 4-point Likert scale intended to measure the presence and severity of urinary, bowel, and sexual function symptoms associated with vaginal prolapse. The KHQ is 21-item questionnaire with a 4-point Likert scale intended to measure the psychosocial effects of urinary incontinence, which include psychological, social, and personal limitations; emotional problems; and sleep or energy disturbance (14). The SPS-Q provides more objective data about symptom prevalence, while the KHQ describes more subjective concerns about those symptoms (15). However, urinary incontinence and prolapse are not expected outcomes of GCS surgery and would probably be considered rare as a consequence of these surgeries by most expert surgeons. For this reason, the utility of using these measures after genital GCS may not have much benefit.

Urinary outcomes in transgender men have largely been measured using function-specific measures not validated for transgender individuals, such as the International Prostate Symptom Score (IPSS) and 24-Hour Voiding Diary (25,26). The IPSS is an 8-question tool reporting outcomes on a 6-point Likert scale, measuring the frequency of urinary symptoms commonly associated with benign prostatic hyperplasia, including retention, frequency, urgency, weak stream, straining, and nocturia, as well as QoL due to urinary symptoms (25). Although this tool may capture some of the symptoms associated with lower urinary tract obstruction secondary to urethral strictures, a common complication of phalloplasty, the IPSS is not validated for this purpose, even in cis-gender males. Moreover, the IPSS will not capture symptoms associated with urinary fistula or symptoms from other common urologic complications of phalloplasty, including persistent vaginal remnant. The 24-Hour Voiding Diary is an exercise that asks individuals to record their urinary activity over a 24-hour period in

order to assess presence, frequency, and severity of urinary symptoms. Information about voiding habits is usually helpful in individuals with urinary problems, however completing the diary is labor intensive.

Sexual health outcomes

A wide variety of PROMs have been used to measure sexual health outcomes in transgender women following vaginoplasty. Most studies use tools validated to measure sexual health in cisgender women (women whose gender identity matches with their sex assigned at birth), unrelated to surgery. These include the Brief Index of Sexual Functioning for Women (BISFW), the Female Sexual Function Index (FSFI) and its adaptation, the Sexual Functioning Index-Gender Spectrum (SFI-GS), Female Genital Self Image Score (FGSIS), Sexual Desire Inventory (SDI), Amsterdam Hyperactive Pelvic Floor Scale-Women (AHPFS-W), and Sheffield Prolapse Symptoms Questionnaire (SPS-Q) (15,16,18,20-22). Of these, the FGSIS, BISFW and SFI-GS are likely to be the most useful, depending on the outcomes to be measured.

The FGSIS may offer an indication of overall sexual satisfaction related to genital anatomy. It is a 7-item scale measuring satisfaction with genitalia globally, and specifically regarding sexual function (20). Although it has not been validated for use in transgender women, it has been modified for use in this population (3). FGSIS adaptation generally reports good outcomes. In one study of 117 vaginoplasty patients, 88% reported being able to get sexually aroused, 83% reported normal genital function, and 83% reported that sexual partners were satisfied with their genitals (35). A smaller study of 15 individuals reported a mean score of 20, near average for cisgender women (20,36).

The BISFW is a 22-item scale measuring sexual interest, activity, and satisfaction (16). It provides data both on details of sexual function and the interpersonal and emotional consequences of sexual health. The BISFW focuses more on sexual activity, satisfaction with sex life, and interpersonal ramifications of sexual health than the FSFIS. It also provides data on a greater number of factors that can contribute to sexual dysfunction, such as dyspareunia, lack of lubrication, anorgasmia, vaginal tightness, incontinence, and vaginal infection, all of which are relevant to transgender women after vaginoplasty. Some of these measurements, such as vaginal lubrication, would be expected to be worse than cisgender women since the most frequent type of vaginoplasty is using a full thickness

skin graft, which does not self-lubricate.

The FSFI only includes measures of lubrication, ability to orgasm, and pain (18). However, the FSFI has been adapted for use in one small sample (n=15) of transgender women, reporting satisfactory functioning (36). The SFI-GS is a newer adaptation of the FSFI for individuals across the gender identity spectrum (19). Five domains are addressed: sexual activity, alone or with a partner; sexual interest; physical arousal, including lubrication, genital engorgement; orgasm; and sexual pain or discomfort. In this 19-item questionnaire, patients are asked to rate their level of interest, arousal, satisfaction or otherwise in the last 6 months. With its gender neutrality, the SFI-GS would be highly applicable for pre- and post-operative comparisons following any form of GCS.

Several partially validated measures have also been utilized. Of these, the Morrison Questionnaire has the benefit of reporting discharge, vaginal odor, and bleeding with intercourse, which most of the other sexual function outcome measures do not (23). It has undergone some limited validation for use in transgender women after vaginoplasty, but it is not completely validated for any purpose. Additionally, some literature uses simple 10-point Likert scales to measure overall sexual satisfaction (37,38). A simple question such as “On a scale of 1–10, how satisfied are you with your genital function?” is broadly informative but does not provide actionable data (37). It will not assess specific aspects of genital functioning, which would be necessary both to guide clinical interventions aimed at improving genital function and satisfaction, and to identify specific shortcomings of these procedures that may be improved.

Sexual health outcomes measures in individuals who have undergone masculinizing genital reconstruction are more limited, potentially due to heterogeneity among surgical outcomes. Metoidioplasties create a smaller phallus that typically does not allow for penetrative intercourse. Pedicled and free flap phalloplasties are subject to variable tactile and erogenous sensation based on surgical technique, and require erectile implants if the patient desires capacity for penetrative intercourse. One study of 49 transgender men after phalloplasty reported sexual outcomes in a semi-validated self-constructed measure (24). This questionnaire was constructed to measure and compare sexual function during masturbation and sexual activity in individuals who did or did not have an erectile prosthesis placed. The questions involved frequency of masturbation before and after surgery, ability to reach orgasm during masturbation

before and after surgery, change in orgasmic feeling after surgery, use of the vagina and clitoris during masturbation prior to surgery, and comparison of sexual arousal before and after surgery. There were no statistically significant differences before and after surgery, or between individuals who did or did not have erectile prosthesis placement.

Most data regarding sexual health and functionality in transgender men after phalloplasty come from *ad-hoc* tools that have undergone no validation process (2). One such study measured patient motivations for phalloplasty or metoidioplasty prior to the procedure and then determined the success of the procedure in meeting these goals (39). The most commonly reported goals were the ability to orgasm (100% desired) and the enabling of sexual intercourse (95%), although these goals were not met by the procedure. The same study found that individuals were, however, more likely to use their genitals for sex after surgery than before (31% versus 78%).

Aesthetic outcomes

Aesthetic outcomes of vaginoplasty have been measured by appropriation of function-specific tools not validated in transgender populations and several tools partially validated for use in GCS (2). The FGSIS, validated in cisgender women, and the Morrison Questionnaire, partially validated for transgender women after GCS, report satisfaction with aesthetic outcomes on a 4- or 5-point Likert scale, reporting generally high satisfaction (20,23). The FGSIS adaptation reported that 80% of 117 patients felt comfortable letting their partner look at their genitals, and the Morrison Questionnaire reported satisfaction of 4.67/5 (23,35). However, neither reports factors that could contribute to aesthetic satisfaction.

Data on aesthetic outcomes of phalloplasty come from the BI-1, as well as non-validated measures (9,26). Outcomes are generally poor. One BI-1 study of 21 phalloplasty patients reported no significant change after the procedure (26). The same study asked 21 phalloplasty patients to rate the ability of phalloplasty to meet their most important pre-operative goals, showing that discomfort in public saunas and showers was not improved to a great extent by the procedure.

Other outcomes

Subjective masculinity and femininity have been measured mostly by non-validated, *ad-hoc* scales. They mostly ask

individuals to rate how masculine or feminine they feel, or how satisfied they are with their masculinity or femininity. One study of 49 transgender women found that individuals felt more feminine after vaginoplasty than before (from 7.6 to 9.01 out of 10, $P < 0.01$) (38). One study of 21 transgender men found no significant change in satisfaction with masculinity following phalloplasty, although it should be noted that they reported feeling very satisfied with their masculinity even before surgery (1.6 out of 5, 1 is most satisfied) (26).

Future directions

In order to understand and meet the needs of TGNB individuals seeking surgical transition, validated PROMs are essential to measure patient experienced QoL, gender dysphoria, and function-specific outcomes following GCS.

From a provider's standpoint, an ideal outcome measure for genital GCS would include: pre- and post-operative assessments of QoL, such as with the SWLS; overall gender dysphoria, and dysphoria specifically related to one's physical appearance, which could be measured with components of the TCS, UGDS-GS and GIDYQ-AA; overall sexual functioning, with the SFI-GS providing a global sexual function assessment for individuals along the gender spectrum, and adaptations of the BISFW and FGSIS for individuals post-vaginoplasty; voiding symptoms, with an adaptation of the IPSS; and post-operative aesthetic satisfaction, with the FGSIS for transwomen. More specific measures of sexual health and aesthetic satisfaction for individuals after phalloplasty and metoidioplasty are needed. These studies reviewed in this article have provided valuable general, sexual, urinary, and aesthetic subjective outcomes data, but have not been validated to do so, limiting the conclusions that can be drawn about individual or relative efficacy of GCS in meeting patient needs.

GENDER-Q shows promise as a new PROM to evaluate gender affirming treatment outcomes (7). Currently in Phase I recruitment, cognitive interviews will be performed to establish content validity for existing BREAST-Q and FACE-Q scales. BREAST-Q and FACE-Q are PROMs developed for breast and facial surgery satisfaction, based upon cisgender populations. Subsequent qualitative interview will elicit new TGNB-specific concepts. The new scales will then be refined and field-tested among an international sample of participants (Phase II). Finally, Phase III studies will involve further psychometric research to determine how well the GENDER-Q scales

measure clinical change after gender affirming treatment. GENDER-Q will ultimately encompass genital GCS as well. To date, there has been no PROM developed through the iterative process of patient and provider focus groups, interviews and field testing in genital reconstructive surgery for TGNB individuals.

Conclusions

PROMs for GCS should be patient-focused, measuring the outcomes that are most important to individuals seeking genital GCS. A number of studies that have collected data regarding goals and preferences of phalloplasty could serve as the first step and these could serve as a basis for question generation (39). QoL outcomes should be sensitive to the TGNB-specific psychosocial factors that influence QoL and gender dysphoria in these individuals (40). Functional outcomes should be sensitive to the unique urinary, sexual, and aesthetic results and potential complications of bottom surgery (4,41). Whichever tools are ultimately developed should be validated not just in transgender women and men, but in gender nonbinary individuals who may seek bottom surgery as well. Such tools will be necessary to study the efficacy of genital GCS surgery and to guide future attempts to improve it.

Acknowledgments

None.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Barone M, Cogliandro A, Di Stefano N, et al. A Systematic Review of Patient-Reported Outcome Measures Following Transsexual Surgery. *Aesthetic Plast Surg* 2017;41:700-13.
2. Andréasson M, Georgas K, Elander A, et al. Patient-Reported Outcome Measures Used in Gender Confirmation Surgery: A Systematic Review. *Plast Reconstr Surg* 2018;141:1026-39.
3. Manrique OJ, Adabi K, Martinez-Jorge J, et al. Complications and Patient-Reported Outcomes in Male-to-Female Vaginoplasty-Where We Are Today: A Systematic Review and Meta-Analysis. *Ann Plast Surg*

- 2018;80:684-91.
4. Frey JD, Poudrier G, Chiodo MV, et al. A Systematic Review of Metoidioplasty and Radial Forearm Flap Phalloplasty in Female-to-male Transgender Genital Reconstruction: Is the "Ideal" Neophallus an Achievable Goal? *Plast Reconstr Surg Glob Open* 2016;4:e1131.
 5. Black N, Jenkinson C. Measuring patients' experiences and outcomes. *BMJ* 2009;339:b2495.
 6. Massie JP, Morrison SD, Smith JR, et al. Patient-Reported Outcomes in Gender Confirming Surgery. *Plast Reconstr Surg* 2017;140:236e-7e.
 7. Klassen AF, Kaur M, Johnson N, et al. International phase I study protocol to develop a patient-reported outcome measure for adolescents and adults receiving gender-affirming treatments (the GENDER-Q). *BMJ Open* 2018;8:e025435.
 8. Diener E, Emmons RA, Larsen RJ, et al. The Satisfaction With Life Scale. *J Pers Assess* 1985;49:71-5.
 9. Lindgren TW, Pauly IB. A body image scale for evaluating transsexuals. *Arch Sex Behav* 1975;4:639-56.
 10. Verschoor AM, Poortinga J. Psychosocial differences between Dutch male and female transsexuals. *Arch Sex Behav* 1988;17:173-8.
 11. Cohen-Kettenis PT, van Goozen SH. Sex reassignment of adolescent transsexuals: a follow-up study. *J Am Acad Child Adolesc Psychiatry* 1997;36:263-71.
 12. McGuire J, Berg DR, Rider N, et al. The Utrecht Gender Dysphoria Scale - Gender Spectrum. 2016. Available online: http://license.umn.edu/technologies/20170216_gender-affirmative-lifespan-approach
 13. Deogracias JJ, Johnson LL, Meyer-Bahlburg HF, et al. The gender identity/gender dysphoria questionnaire for adolescents and adults. *J Sex Res* 2007;44:370-9.
 14. Kelleher CJ, Cardozo LD, Khullar V, et al. A new questionnaire to assess the quality of life of urinary incontinent women. *Br J Obstet Gynaecol* 1997;104:1374-9.
 15. Bradshaw HD, Hiller L, Farkas AG, et al. Development and psychometric testing of a symptom index for pelvic organ prolapse. *J Obstet Gynaecol* 2006;26:241-52.
 16. Taylor JE, Rosen RC, Leiblum SR. Self-report assessment of female sexual function: psychometric evaluation of the Brief Index of Sexual Functioning for Women. *Arch Sex Behav* 1994;23:627-43.
 17. Reynolds CF 3rd, Frank E, Thase ME, et al. Assessment of sexual function in depressed, impotent, and healthy men: factor analysis of a Brief Sexual Function Questionnaire for men. *Psychiatry Res* 1988;24:231-50.
 18. Rosen R, Brown C, Heiman J, et al. The Female Sexual Function Index (FSFI): a multidimensional self-report instrument for the assessment of female sexual function. *J Sex Marital Ther* 2000;26:191-208.
 19. Spencer KD, Vencill J. Sexual Functioning Index-Gender Spectrum. University of Minnesota; 2017. Available online: <https://www.sexualhealth.umn.edu/ncgsh/measures>
 20. Herbenick D, Reece M. Development and validation of the female genital self-image scale. *J Sex Med* 2010;7:1822-30.
 21. Spector IP, Carey MP, Steinberg L. The sexual desire inventory: development, factor structure, and evidence of reliability. *J Sex Marital Ther* 1996;22:175-90.
 22. Postma R, Bicanic I, van der Vaart H, et al. Pelvic floor muscle problems mediate sexual problems in young adult rape victims. *J Sex Med* 2013;10:1978-87.
 23. Morrison SD, Satterwhite T, Grant DW, et al. Long-Term Outcomes of Rectosigmoid Neocolporrhaphy in Male-to-Female Gender Reassignment Surgery. *Plast Reconstr Surg* 2015;136:386-94.
 24. Wierckx K, Van Caenegem E, Elaut E, et al. Quality of life and sexual health after sex reassignment surgery in transsexual men. *J Sex Med* 2011;8:3379-88.
 25. Barry MJ, Fowler FJ Jr, O'leary MP, et al. The American Urological Association Symptom Index for Benign Prostatic Hyperplasia. *J Urol* 2017;197:S189-97.
 26. van de Grift TC, Pigot GLS, Boudhan S, et al. A Longitudinal Study of Motivations Before and Psychosexual Outcomes After Genital Gender-Confirming Surgery in Transmen. *J Sex Med* 2017;14:1621-8.
 27. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. *Health Econ* 1993;2:217-27.
 28. Cardoso da Silva D, Schwarz K, Fontanari AM, et al. WHOQOL-100 Before and After Sex Reassignment Surgery in Brazilian Male-to-Female Transsexual Individuals. *J Sex Med* 2016;13:988-93.
 29. James SE, Herman JL, Rankin S, et al. The Report of the 2015 U.S. Transgender Survey. Washington, DC: National Center for Transgender Equality, 2016.
 30. Diagnostic and statistical manual of mental disorders: DSM-5. 5th edition. Arlington, VA: American Psychiatric Publishing, 2013.
 31. Kozee HB, Tylka TL, Bauerband LA. Measuring Transgender Individuals' Comfort With Gender Identity and Appearance: Development and Validation of the Transgender Congruence Scale. *Psychol Women Q* 2012;36:179-96.
 32. Eisenberg ME, Gower AL, McMorris BJ, et al. Risk and Protective Factors in the Lives of Transgender/Gender Nonconforming Adolescents. *J Adolesc Health*

- 2017;61:521-6.
33. White Hughto JM, Reisner SL. A Systematic Review of the Effects of Hormone Therapy on Psychological Functioning and Quality of Life in Transgender Individuals. *Transgend Health* 2016;1:21-31.
 34. Chaiet SR, Morrison SD, Streed CG Jr. Gender Confirmation Surgery and Terminology in Transgender Health. *JAMA Surg* 2017;152:1089-90.
 35. Massie JP, Morrison SD, Van Maasdam J, et al. Predictors of Patient Satisfaction and Postoperative Complications in Penile Inversion Vaginoplasty. *Plast Reconstr Surg* 2018;141:911e-21e.
 36. Manrique OJ, Sabbagh MD, Ciudad P, et al. Gender-Confirmation Surgery Using the Pedicle Transverse Colon Flap for Vaginal Reconstruction: A Clinical Outcome and Sexual Function Evaluation Study. *Plast Reconstr Surg* 2018;141:767-71.
 37. Lawrence AA. Factors associated with satisfaction or regret following male-to-female sex reassignment surgery. *Arch Sex Behav* 2003;32:299-315.
 38. Zavlin D, Schaff J, Lelle JD, et al. Male-to-Female Sex Reassignment Surgery using the Combined Vaginoplasty Technique: Satisfaction of Transgender Patients with Aesthetic, Functional, and Sexual Outcomes. *Aesthetic Plast Surg* 2018;42:178-87.
 39. Jacobsson J, Andreasson M, Kolby L, et al. Patients' Priorities Regarding Female-to-Male Gender Affirmation Surgery of the Genitalia-A Pilot Study of 47 Patients in Sweden. *J Sex Med* 2017;14:857-64.
 40. Bariola E, Lyons A, Leonard W, et al. Demographic and Psychosocial Factors Associated With Psychological Distress and Resilience Among Transgender Individuals. *Am J Public Health* 2015;105:2108-16.
 41. Dy GW, Sun J, Granieri MA, et al. Reconstructive Management Pearls for the Transgender Patient. *Curr Urol Rep* 2018;19:36.

Cite this article as: Dy GW, Nolan IT, Hotaling J, Myers JB. Patient reported outcome measures and quality of life assessment in genital gender confirming surgery. *Transl Androl Urol* 2019;8(3):228-240. doi: 10.21037/tau.2019.05.04