OXFORD

Development of Health-Related Quality of Life Instruments for Young Children With Disorders of Sex Development (DSD) and Their Parents

Adrianne N. Alpern,[1] PhD, Melissa Gardner,[2] MA, Barry Kogan,[3] MD, David E. Sandberg,[2] PhD, and Alexandra L. Quittner,[1] PhD

[1]Department of Psychology, University of Miami, [2]Department of Pediatrics; Division of Pediatric Psychology and Child Health Evaluation & Research (CHEAR) Unit, University of Michigan, and [3]Division of Urology, Albany Medical College

For questions regarding the development of the QOL-DSD-Proxy and QOL-DSD-Parent (e.g., psychometric properties, strategies for translation) or to obtain the instruments, contact Alexandra L. Quittner (aquittner@miami.edu). Inquiries regarding the questionnaires' use in the clinical or clinical research environment should be directed to David E. Sandberg (dsandber@med.umich.edu).

Abstract

Objective Research in disorders of sex development (DSD) is hindered by a lack of standardized measures sensitive to the experiences of affected children and families. We developed and evaluated parent proxy (children 2–6 years) and parent self-report (children ≤6 years) health-related quality of life (HRQoL) instruments for DSD. **Methods** Items were derived from focus groups and open-ended interviews. Clarity and comprehensiveness were assessed with cognitive interviews. Psychometric properties were examined in a field survey of 94 families. **Results** Measures demonstrated adequate to good psychometrics, including internal consistency, test–retest reliability, convergent validity, and ability to detect known-group differences. Parents reported greatest stress on *Early Experiences*, *Surgery*, and *Future Concerns* scales. **Conclusions** These instruments identify patients' and families' needs, monitor health and quality of life status, and can evaluate clinical interventions. Findings highlight the need for improved psychosocial support during the diagnostic period, better parent–provider communication, and shared decision-making. HRQoL measures are needed for older youth.

Key words: disorders of sex development; health-related quality of life; patient-reported outcome; psychometric evaluation.

Introduction

Disorders (or differences) of sex development (DSD), formerly referred to as hermaphroditism or intersex, are "congenital conditions in which development of chromosomal, gonadal, or anatomic sex is atypical" (Lee, Houk, Ahmed, Hughes, & Participants in the International Consensus Conference on Intersex, 2006). Although this umbrella term encompasses diverse conditions affecting sex determination and/or sex differentiation (Achermann & Hughes, 2011), DSD share physical, social, and emotional sequelae. For example, gender announcement in newborns may

be delayed owing to atypical genitalia or a mismatch between genetic sex, gonadal function, and/or anatomic sex. Children can experience difficulties with toileting or become self-conscious about genital appearance. Parents may question the gender assignment decision, particularly when the child displays gender-atypical behavior. They may also worry about stigma and attempt to conceal the child's condition from others. Moreover, parents must make difficult decisions regarding surgery and, in select DSD, they are burdened with managing chronic medications (Sandberg, Gardner, & Cohen-Kettenis, 2012). In part

owing to low prevalence rates (estimated incidence of 1 in 4,500 live births; Hughes, Houk, Ahmed, & Lee, 2006), little research exists describing how these conditions affect the daily functioning of children with DSD and their caregivers (Sandberg et al., 2012). The purpose of this study was to develop two DSD-specific health-related quality of life (HRQoL) measures: a parent proxy measure for children aged 2–6 years and a self-report measure for parents of children aged birth to 6 years.

The Consensus Statement (Lee et al., 2006) classifies discrete DSD diagnoses into three categories based on karyotype: sex chromosome DSD; 46,XY DSD; and 46,XX DSD. The category of sex chromosome DSD includes the diagnosis of mixed gonadal dysgenesis in which there is more than one cell line in the body (45,X and 46,XY) and is associated with genitalia ranging from ambiguous to typical male or female. An example of 46,XY DSD is proximal hypospadias, which refers to a condition in which the opening of the urethra is not on the tip of the penis; it can be located on the underside of the penis near the tip (distal hypospadias) or along the shaft, at the base of the penis, or within the scrotum (proximal hypospadias). Proximal hypospadias is associated with abnormal development of the underside side of the penis and is uniformly associated with a marked downward curvature of the penis when erect (i.e., chordee). Finally, the classic form of congenital adrenal hyperplasia (CAH) in girls represents one example of a 46,XX DSD. In CAH, the adrenal glands do not produce enough corticosteroids, instead overproducing androgens; in girls, this results in masculinization of the internal and external genital structures. External genital appearance can range from an enlarged clitoris to a fully formed penis.

Importantly, both research and clinical care in DSD have been hindered by a lack of standardized measures that are sensitive to the unique experiences of this population (Berenbaum, Korman Bryk, Duck, & Resnick, 2004; Migeon et al., 2002; Sandberg & Mazur, 2014; Schober et al., 2012; Wisniewski et al., 2001). Recent studies have used measures of psychological distress (e.g., anxiety, depression), generic HRQoL questionnaires (Krupp et al., 2014; Pasterski, Mastroyannopoulou, Wright, Zucker, & Hughes, 2014; Wolfe-Christensen, Fedele, Mullins, Lakshmanan, & Wisniewski, 2014), or ad hoc scales that have not been validated (Lux, Kropf, Kleinemeier, Jurgensen, & Thyen, 2009). Most studies narrowly focused on adult surgical and sexual function outcomes, or psychosexual differentiation (i.e., gender identity, gender role, and sexual orientation; Lee et al., 2012; Pasterski et al., 2015; Schönbucher et al., 2012; Stout, Litvak, Robbins, & Sandberg, 2010; van der Zwan et al., 2013).

Sensitive outcome measures are particularly important because clinical management of individuals with DSD has been controversial (e.g., desirability and timing of genital or gonadal surgery). For example, negative surgical outcomes voiced by former patients have influenced attitudes about clinical practice, including recommendations about if or when to perform surgery (Diamond & Garland, 2014; Feder, 2014; Karkazis, 2008; Lantos, 2013). The lack of DSD-specific assessment tools makes it challenging to systematically inform changes in practice and evaluate outcomes.

HRQoL tools assess the effects of a medical condition on physical and psychosocial functioning (e.g., social, emotional, role-related), and have been used to weigh the costs and benefits of healthcare interventions, monitor outcomes, and aid in clinical decision-making (Quittner, Davis, & Modi, 2003; Spilker, 1996); however, they are rarely used in DSD care or research (Schober et al., 2012). Accordingly, the *Strategic Plan for Pediatric Urology Research* prioritized assessment of the clinical outcomes of children with DSD and encouraged the development of sensitive HRQoL measures (National Institute of Diabetes and Digestive and Kidney Diseases, 2006). Such tools will enable clinicians to: (1) identify specific areas of need for patients and their families, (2) measure the effects of medical and surgical interventions on caregiver and family functioning, and (3) provide empirical evidence to guide patient management (Guyatt, Feeny, & Patrick, 1993; National Institute of Diabetes and Digestive and Kidney Diseases, 2006; Spilker, 1996).

Finally, little is known about the HRQoL of children or their caregivers soon after detection of the DSD, when decisions are made about the child's care. In a society that views biological sex as a binary, it may be difficult for parents to understand that their child's biological sex does not conform to "male" or "female." Parents are often unaware that gender identity, chromosomes, and anatomy are not always concordant, and they may be uncertain about their child's future gender identity. Following detection of the DSD, parents must learn new, complex medical information, consider the medical and psychological implications of their decisions, cope with ongoing medical tests and procedures (potential for multiple operations), handle financial burdens, and face strains on family roles and relationships (Kogan et al., 2012; Sandberg et al., 2012; Sanders, Carter, & Goodacre, 2011; Sanders, Carter, & Goodacre, 2012). There are currently no standardized measures that capture these parental stressors.

The goal of this study was to develop and validate two HRQoL instruments for young children with DSD and their caregivers: a proxy measure of HRQoL in parents of young children with DSD (age 2 to 6

years) and a self-report measure for parents. Measures were developed in accordance with the Food and Drug Administration's guidance on patient-reported outcomes (Food and Drug Administration, 2009) and consistent with established criteria for measure development (Holmbeck & Devine, 2009) using a relatively large, medical-chart-selected sample of children with DSD. In terms of convergent validity with established measures, it was hypothesized that DSD-specific domains on the parent proxy report would be correlated with like domains on the PedsQL (e.g., Physical, Social, and Emotional Functioning; Varni, Seid, & Kurtin, 2001; Varni, Seid, & Rode, 1999). On the parent self-report measure, it was hypothesized that genital atypicality and number of surgical procedures would be associated with Surgery and Earliest Experiences; the Hopkins Symptom Checklist (HSCL; Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974) would be correlated with Social and Emotional Functioning; select subscales of the Parenting Stress Index (PSI; Abidin, 1995) would be associated with Social Functioning, Role Functioning, Emotional Functioning, Medications, and Doctor's Visits; The Experiences and Reactions Questionnaire (ERQ; Rolston, Gardner, Vilain, & Sandberg, 2015; Sandberg, Gardner, & Rolston, 2010) would be correlated with Decision-Making, Future Concerns, Healthcare Communication and Information, and

Disclosure; the Impact on Family Scale (IoF) would be associated with Gender Expression, Social Functioning, Future Concerns, and Medications; and the Decisional Regret Scale (DRS; Brehaut et al., 2003), Decisional Conflict Scale (DCS; O'Connor, 1995), and Trust in Physical Scale (Moseley, Clark, Gebremariam, Sternthal, & Kemper, 2006) would be correlated with select domains related to Decision-Making, Healthcare Communication, Surgery, and Doctor's Visits.

Methods

The steps of measure development as implemented in this study are described in Figure 1.

Participants

Focus Groups

Expert clinicians (pediatric urologists, $n = 7$; pediatric endocrinologists, $n = 10$; mental health professionals, $n = 4$), DSD patient advocates ($n = 4$), and parents of DSD-affected children (newborn to 6 years; $n = 11$) participated in focus groups and interviews to establish initial questionnaire content. A description of the focus group methodology and results are published elsewhere (Kogan et al., 2012). A brief overview of the measure development process is highlighted in Figure 1.
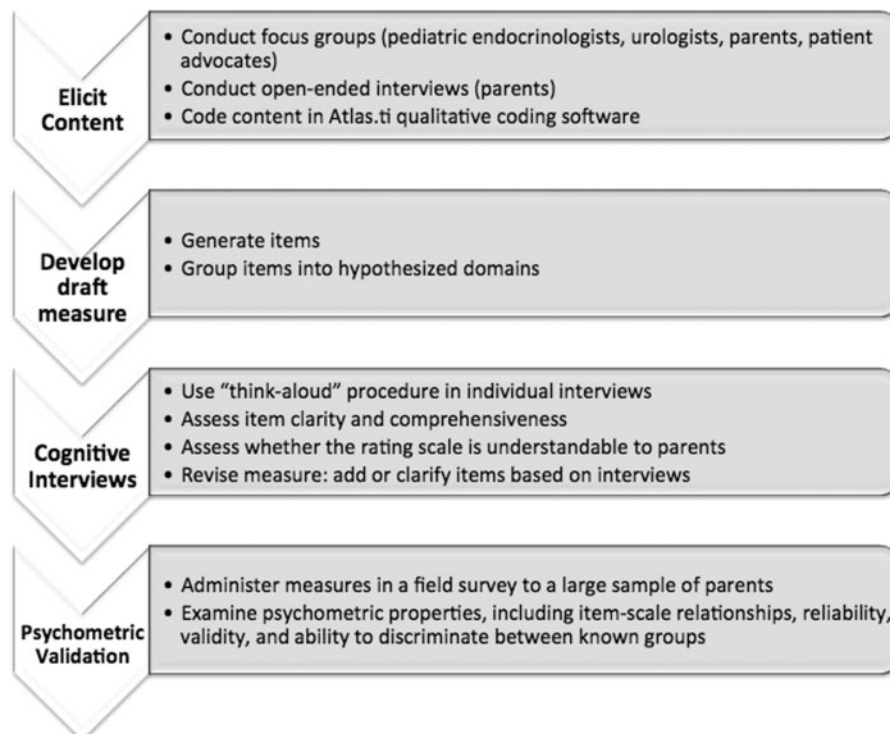

Elicit Content
- Conduct focus groups (pediatric endocrinologists, urologists, parents, patient advocates)
- Conduct open-ended interviews (parents)
- Code content in Atlas.ti qualitative coding software

Develop draft measure
- Generate items
- Group items into hypothesized domains

Cognitive Interviews
- Use "think-aloud" procedure in individual interviews
- Assess item clarity and comprehensiveness
- Assess whether the rating scale is understandable to parents
- Revise measure: add or clarify items based on interviews

Psychometric Validation
- Administer measures in a field survey to a large sample of parents
- Examine psychometric properties, including item-scale relationships, reliability, validity, and ability to discriminate between known groups

Figure 1. Steps of measure development. These steps are consistent with the Food and Drug Administration's regulatory guidance on the development of patient-reported outcomes (FDA, 2009).

Cognitive Interviews

Forty-two parents of 28 children (61% 46,XY DSD; 25% 46,XX DSD; 14% sex chromosome DSD) participated in cognitive interviews, in which they completed preliminary versions of both instruments and were asked to "think-aloud" as they responded (Table I).

Participants were recruited if they had a child with a DSD (birth to 6.75 years at the time of recruitment) and received care at any of four participating academic medical centers. To reduce selection bias, cases were chart-selected based on ICD-9 diagnostic codes, followed by chart review to ensure eligibility. Because of the low prevalence, children with sex chromosome DSD (e.g., 45,X/46,XY mixed gonadal dysgenesis) were deliberately over-sampled to ensure representation in the study sample. Children with significant developmental delays or physical disabilities documented in the medical record (e.g., cloacal exstrophy) were excluded because these features could potentially confound interpretation of the findings. Those diagnosed with Klinefelter (47,XXY or its variants) and Turner syndrome (45,X or its variants) were also excluded, in addition to those born with isolated distal hypospadias, because these conditions are not uniformly classified as DSD (Wit, Ranke, & Kelnar, 2007). Parental eligibility was restricted to those whose primary language was English.

Field Survey

A total of 130 parents/caregivers of 94 affected children participated in the field survey. Families were recruited at 12 medical centers to obtain an adequate sample size for the range of DSD conditions, many of which are classified as rare diseases, as well as to sample from various geographic regions, using the same methodology used in the cognitive interview phase. The majority of children ($n = 59$) were diagnosed with 46,XY DSD, followed by ($n = 31$) 46,XX DSD, and ($n = 4$) sex chromosome DSD (Table II). A 56% participation rate was achieved (Table III). Institutional review board (IRB) approval and parental informed consent were obtained at all sites for all phases of measure development. Because of the sensitive nature of DSD and families' desire for added privacy, the IRB did not permit retention of demographic information for families that declined to participate; thus, it was not possible to examine demographic differences between families that agreed to participate and those that declined.

Measures

Demographic and Medical Information

Parents reported their age, race, ethnicity, education, and employment status, as well as their child's race and ethnicity. Medical data were extracted from charts, including the child's date of birth, gender (i.e., gender announced at birth, assignment after a delay, or reassignment), diagnosis, number of genital operations, age at time of operations, and post-surgical genital appearance and function. Genital atypicality (defined as atypicality of genitals relative to gender of rearing before reconstructive surgery) was assessed based on descriptions documented in the medical record. The Quigley scale was utilized for children reared as boys (Quigley et al., 1995); scores ranged from 1 to 7, with 7 representing a female-typical genital appearance. The Prader scale was utilized for children reared as girls (Achermann & Hughes, 2011); scores ranged from 0 to 5, with 5 representing a male-typical genital appearance. Scores were standardized to a 0–100 scale, with higher scores indicating greater atypicality.

The Quality of Life DSD Proxy Report (QOL-DSD-Proxy) assesses the HRQoL of children with DSD aged 2–6 years using parents as proxy respondents. Items were generated from focus groups and open-ended interviews with adult patients, parents, advocates, and healthcare providers (e.g., endocrinologists, urologists, psychologists; Kogan et al., 2012). Interviews were audiotaped, transcribed, and coded in Atlas.ti 6.0 (Muhr & Freise, 2004) to identify themes and ensure saturation of content. Next, items reflecting the most frequently endorsed and relevant issues were identified and used for item generation. Cognitive testing, using a "think aloud" procedure (Schwarz & Sudman, 1996; Sudman, Bradburn, & Schwarz, 1996), was performed on the preliminary set of items (Version 1.0) to assess clarity and comprehensiveness. Follow-up probes by the interviewer queried their understanding of the questions and answer

Table I. Cognitive Interview Participant Demographics

Participant characteristics	Parent report	
	Proxy	Self
Index cases (*n*)	23[a]	28[b]
Parents (*n*)	34	42
Mother	61.8%	64.3%
Father	35.3%	33.3%
Other (e.g., grandparent)	2.9%	2.4%
Parent race/ethnicity		
White	64.7%	69.1%
Hispanic origin	17.6%	14.3%
African American/Black	8.8%	7.1%
Other	5.9%	4.9%
Did not answer	5.9%	7.1%
Child age, years (mean, *SD*)	4.92 (1.55)	4.39 (1.94)
Child DSD category (*n*)	23	28
Sex chromosome	8.7%	14.3%
46,XY	69.6%	60.7%
46,XX	21.7%	25.0%

[a]Parents of children aged 2–6 years.
[b]Parents of children from birth to 6 years.

Table II. Field Study Participant Demographics

Participant characteristics	Parent report	
	Proxy	Self
Index cases (n)	76[a]	94[b]
Parents (n)	103	130
Mother	67.0%	65.4%
Father	33.0%	34.6%
Parent race		
White	69.7%	71.3%
Asian or Pacific Islander	7.9%	9.6%
African American	3.9%	3.2%
American Indian or Alaska Native	1.3%	1.1%
Middle Eastern	1.3%	1.1%
Other or Mixed race	14.5%	12.8%
Did not answer	1.3%	1.1%
Parent ethnicity: Hispanic	13.2%	11.7%
Parent education		
Less than high school	2.6%	3.2%
High school or equivalent	11.8%	16.0%
Some college	22.4%	22.3%
Bachelor's degree or equivalent	31.6%	26.6%
Graduate or professional degree	30.3%	30.9%
Did not answer	1.3%	1.1%
Child gender of rearing		
Boy	57.9%	56.4%
Girl	42.1%	42.6%
Unassigned or did not answer	0%	1.1%
Child age, years (mean, SD)	4.54 (1.46)	4.02 (1.77)
Child DSD category & diagnosis		
Sex chromosome (n, %)	4 (5.3%)	4 (4.3%)
46,XY (n)	45	59
Disorder of gonadal (testicular) development	6.6%	7.4%
Disorder of androgen synthesis or action	10.5%	12.8%
Other 46,XY DSD (e.g., proximal hypospadias)	42.1%	42.6%
46,XX (n)	27	31
Disorder of gonadal (ovarian) development	1.3%	1.1%
Androgen excess	30.3%	27.7%
Other 46,XX DSD	3.9%	4.3%
Genital atypicality rating (0–100) (mean, SD)	37.71 (20.46)	37.58 (20.56)
Number of surgical procedures		
Mean (SD)	2.15 (1.72)	1.96 (1.62)
Range	0–8	0–8

[a]Parents of children aged 2–6 years.
[b]Parents of children from birth to 6 years.

Table III. Field Study Participation Rate

Participant flow	n (%)
Total eligible index cases	247
Lost to follow-up	78 (31.6%)
Declined to participate	44 (17.8%)
No response	35 (14.2%)
Valid index cases (eligible minus lost)	169
Index cases	94 (55.6%)
Participating caregivers	130

choices. Finally, for each content area (e.g., Physician Functioning, Gender Expression, Social Functioning), respondents were asked if any important content was missing.

The QOL-DSD-Proxy was revised based on these results, leading to Version 2.0, which was used in the validation study. It consisted of 25 items across five domains: Physical Functioning, Gender Expression, Social Functioning, Emotional Functioning, and Medical Care. Using a 2-week recall period for the first four domains, respondents indicated the extent to which each item was true for *their child* on a 5-point Likert scale. Items in the Medical Care domain required respondents to indicate when their child's last doctor's visit occurred, and to recall their child's functioning at that time. Responses were standardized from 0 to 100 to facilitate interpretation, with higher scores representing better HRQoL.

The QOL-DSD-Parent assesses parents' HRQoL in relation to their children with DSD (age birth to 6 years). The development process was identical to that described earlier and the sample was the same, with the addition of 18 parents of children under age 2 (Table II). The QOL-DSD-Parent v.2.0 consisted of 67 items in 12 domains: Decision-Making, Role Functioning and Family Activities, Gender Expression, Social Functioning, Emotional Functioning, Future Concerns, Healthcare Communication and Information, Disclosure, Medications, Surgery, Doctor's Visits, and Earliest Experiences (e.g., stress waiting for a diagnosis). Similar to the proxy report, parents rated their own functioning on a 5-point Likert scale, using a 2-week recall. Responses were standardized from 0 to 100, with higher scores indicating better HRQoL. Parents were also asked to comment on: "What concerns you most about your child at this time?"

A comprehensive set of measures were used to evaluate convergent validity; details (content, psychometrics) are summarized in Table IV.

Statistical Analyses

Item- and scale-level analyses are presented on the QOL-DSD-Proxy first, followed by the QOL-DSD-Parent. Given the sample size guidelines for hierarchical modeling with dyadic data, too few index cases were available to conduct item- and scale-level psychometric analyses dyadically to account for dependencies in the data (i.e., using both a mother and father reporting on the same child; Bell, Ferron, & Kromrey, 2008). We considered randomly selecting one parent per child for inclusion in these analyses; however, given the systematic differences between mothers' and fathers' HRQoL ratings, this might have produced distorted results (Langberg et al., 2010; Schroeder, Hood, & Hughes, 2010). Thus, data from mothers were selected for psychometric analysis because more mothers responded, and fathers' descriptive data were also examined. Informant discrepancies and scale invariance were beyond the scope of this paper and will be addressed in a future publication.

Item-Level Analyses

To test the fit between items and their hypothesized scales, multitrait analyses were conducted for each QOL-DSD measure. This type of analysis was developed for smaller samples and evaluates the correlation between items and their hypothesized versus competing scales (adjusted for overlap; Hays, Anderson, & Revicki, 1993; Hays & Hayashi, 1990). Two criteria were used to retain items: (1) item-total correlations should be 0.40 or higher with the intended scale, and (2) item-total correlations should be higher for the intended versus competing scale. If an item was significantly correlated with multiple scales,

the correlation coefficient for the competing scale could be no more than one standard error greater than the correlation with the intended scale (Hays & Hayashi, 1990).

Scale-Level Analyses

The percentage of scores falling at the extreme ends of the scaling range was examined (McHorney, Ware, & Raczek, 1993), and floor and ceiling effects (i.e., % of respondents with a score of 0 or 100) were calculated for each domain.

Two types of reliability were calculated: (1) internal consistency using Cronbach's alpha, and (2) test–retest reliability (Cronbach, 1951; Nunnally & Berstein, 1994). Reliabilities of 0.70 or greater are considered adequate for comparing patient groups. Test–retest reliability was examined in a subsample of mothers ($n = 25$) who completed the measure twice, 10–14 days apart. Intraclass correlation coefficients (ICCs) were computed with a target of $>.50$ (Table V).

Convergent validity was examined using Pearson correlations with a generic HRQoL scale (PedsQL^{TM}; Varni et al., 1999, 2001) and a measure of genital atypicality. We also examined convergence between the QOL-DSD-Proxy and the Teacher Report Form (TRF; Achenbach & Rescorla, 2001) for children 1.5 years or older with available teacher data ($n = 17$). For the QOL-DSD-Parent, we computed correlations between the instrument and selected scales of the PSI (Abidin, 1995), IoF (Stein & Riessman, 1980), HSCL (Derogatis et al., 1974), DCS (O'Connor, 1995), DRS (Brehaut et al., 2003), Trust in Physician Scale (TiPS; Moseley et al., 2006), ERQ (Sandberg et al., 2010; Rolston et al., 2015), and genital atypicality (described above; Table IV). We matched relevant scales to the criterion measure using a priori decisions about construct overlap (e.g., correlating the DCS and DRS with the QOL-DSD-Parent Decision-Making scale, shown in bold in Table VI).

Known-group differences were evaluated to determine whether both versions of the QOL-DSD could discriminate between children above or below the median genital atypicality rating (≥ 50 vs. <50), and between those with or without uncertainty regarding gender assignment (i.e., gender announced at birth and not changed vs. gender assignment delayed or changed at any point). One-way analyses of variance were conducted to examine group differences. For all scales with skewed distributions (skew >2; QOL-DSD-Parent, Gender Expression scale only), nonparametric Mann–Whitney U tests were conducted.

Results

For the QOL-DSD-Proxy, data were available for 73 mothers and 30 fathers; for the QOL-DSD-Parent,

Table IV. Measures Utilized in Field Study to Assess Convergent Validity

Measure	Scales	Higher scores represent	Internal consistency		Support for validity
			Prior research	Study sample	
Parenting Stress Index (PSI)	Role Restriction; Sense of Competence	More parenting stress	>0.70	0.79–0.84	Strong
Impact on Family Scale (IoF)	General Negative Impact	Better parental functioning/less distress	>0.80	0.94	Strong
Pediatric Quality of Life Inventory™ - Toddler or Young Child Form (PedsQL)	Physical; Emotional; Social; School Functioning	Better child HRQoL	0.86–0.90	Toddler: 0.76–0.86 Young Child: 0.73–0.92	Strong
Hopkins Symptom Checklist (HSCL)	Somatization, Obsessive-Compulsive Symptoms; Interpersonal Sensitivity; Depression; Anxiety; Total Symptom Score	More symptoms/distress experienced by the parent	0.85–0.87	0.79–0.90	Strong
Decisional Conflict Scale (DCS)	Perceived Effective Decision Making (Past, Current, & Total Score)	Greater (worse) parental decisional conflict	0.78–0.92	0.52–0.85	Good
Decisional Regret Scale (DRS)	Total Score	More parental regret and dysfunction	0.81–0.92	0.90	Strong
Trust in Physician Scale (TiPS)	Total score	Greater trust and confidence in the child's healthcare provider	0.85–0.90	0.91	Strong
Experiences and Reactions Questionnaire (ERQ)	Others' Reactions; Personal Reactions; Total Score	Poorer/worse reactions and perceived stigma	0.83	0.78–0.79	Strong
Teacher Report Form (TRF)[a]	Emotionally Reactive; Internalizing Problems; Externalizing Problems; Somatic Complaints; Total Problems; Adaptability	Greater child pathology Better adaptive functioning	0.72–0.97	0.81–0.95	Strong

[a]All measures were completed by parents, except for the TRF which was completed by the child's teacher or daycare provider.

Table V. Descriptive Statistics, Internal Consistency, and Test–Retest Data for the QOL-DSD-Proxy and QOL-DSD-Parent

QOL-DSD scales	n^a	Minimum	Maximum	Mean	SD	% Floor	% Ceiling	Cronbach's α	Test–retest ICC
Parent proxy report: mothers									
Physical Functioning	73	30.00	95.00	83.06	15.35	0.00	0.00	.53	.89
Gender Expression	73	0.00	100.00	83.19	20.28	1.37	32.88	.72	.65
Socio-Emotional Functioning	73	46.88	100.00	82.74	13.46	0.00	8.22	.85	.90
Medical Concerns	71	0.00	100.00	64.98	24.49	1.41	5.63	.82	.85
Parent self-report: mothers									
Decision-Making	89	41.67	100.00	87.03	14.02	0.00	34.83	.81	.85
Role Functioning	85	0.00	100.00	75.39	28.98	4.71	28.24	.86	.97
Gender Expression	88	37.50	100.00	94.13	13.09	0.00	70.45	.82	.11
Social Functioning	89	8.33	100.00	86.89	21.06	0.00	55.06	.72	.74
Emotional Functioning	89	16.67	95.83	65.73	16.12	0.00	0.00	.81	.94
Future Concerns	89	0.00	100.00	55.37	25.86	1.12	1.12	.88	.95
Communication and Information	87	25.00	100.00	74.86	16.93	0.00	6.90	.70	.92
Disclosure	88	12.50	100.00	64.03	24.52	0.00	12.50	.81	.88
Medications	33	0.00	100.00	70.39	28.20	3.03	21.21	.75	.75 ($n=6$)
Surgery	78	0.00	100.00	38.03	25.12	11.54	1.28	.61	.88
Doctor's Visits	87	0.00	100.00	72.08	27.81	2.30	26.44	.78	.61
Early Experiences	89	0.00	93.75	25.91	26.46	28.09	1.12	.80	.81
Parent proxy report: fathers									
Physical Functioning	30	50.00	95.00	85.50	12.29	0.00	0.00	n/a	n/a
Gender Expression	30	37.50	100.00	86.04	18.18	0.00	40.00	n/a	n/a
Socio-Emotional Functioning	30	33.33	100.00	85.88	13.90	0.00	13.33	n/a	n/a
Medical Concerns	29	15.00	100.00	78.10	22.56	0.00	20.69	n/a	n/a
Parent self-report: fathers	41	50.00	100.00	86.69	13.83	0.00	34.15	n/a	n/a
Decision-Making	40	25.00	100.00	85.89	21.08	0.00	47.50	n/a	n/a
Role Functioning	41	56.25	100.00	96.14	8.67	0.00	73.17	n/a	n/a
Gender Expression	41	66.67	100.00	93.50	10.46	0.00	63.41	n/a	n/a
Social Functioning	41	41.67	100.00	74.90	13.23	0.00	2.44	n/a	n/a
Emotional Functioning	41	18.75	100.00	64.79	21.89	0.00	2.44	n/a	n/a
Future Concerns	40	42.86	100.00	79.14	13.71	0.00	7.50	n/a	n/a
Communication and Information	39	10.00	100.00	69.97	23.15	0.00	17.95	n/a	n/a
Disclosure	16	50.00	100.00	85.55	16.80	0.00	31.25	n/a	n/a
Medications	37	0.00	100.00	49.10	28.31	8.11	5.41	n/a	n/a
Surgery	41	0.00	100.00	81.10	24.51	4.88	41.46	n/a	n/a
Doctor's Visits	41	0.00	100.00	33.38	25.80	17.07	2.44	n/a	n/a
Early Experiences	30	50.00	95.00	85.50	12.29	0.00	0.00	n/a	n/a

[a]n's vary by scale because not all children have had surgery or take medication.
[b]Test–retest data were available for $n=25$ for parent self-report, $n=21$ for parent proxy-report, unless otherwise noted.

Table VI. QOL-DSD-Proxy Convergent Validity

		QOL-DSD-Proxy Scale							
		Physical Functioning		Gender Expression		Socio-Emotional Functioning		Medications	
Validation measure	Item/Scale	n	r	n	r	n	r	n	r
Medical chart	Child Age (years)	73	0.32**	73	0.13	73	−0.03	71	0.07
	Genital Atypicality	**71**	**0.18**	71	0.07	71	−0.02	69	0.00
	n surgeries	**62**	**−0.20**	62	−0.05	62	−0.06	61	0.01
PedsQL™ Toddler	Physical Functioning	**44**	**0.01**	44	0.13	44	0.58**	43	0.17
	Emotional Functioning	44	0.20	44	0.31*	**44**	**0.66****	43	0.32*
	Social Functioning	43	−0.01	43	0.19	**43**	**0.70****	42	0.10
	School Functioning	33	0.33	33	0.06	**33**	**0.67****	32	0.24
PedsQL™ Young Child	Physical Functioning	**27**	**0.16**	27	0.62**	27	0.63**	26	0.30
	Emotional Functioning	27	0.24	27	−0.03	**27**	**0.45***	26	0.41*
	Social Functioning	27	0.44*	27	0.20	**27**	**0.75****	26	0.68**
	School Functioning	26	0.49*	26	0.42*	**26**	**0.64****	25	0.47*
Teacher Report Form (TRF)	Emotionally Reactive	17	0.06	17	−0.01	**17**	**−0.46**	16	−0.04
	Somatic Problems	**17**	**0.12**	17	−0.20	17	−0.42	**16**	**−0.41**
	Internalizing Problems	17	0.12	17	−0.20	**17**	**−0.42**	16	−0.41
	Externalizing Problems	17	0.11	17	−0.21	**17**	**−0.41**	16	−0.41
	Total Problems	17	0.11	17	−0.21	**17**	**−0.41**	16	−0.41
	Adaptability	17	−0.07	17	0.02	**17**	**0.46**	16	0.04

Note. $r =$ Pearson's r. Hypothesized relationships between scales are shown in bold.
*$p < .05$, **$p < .01$.

data were available for 89 mothers and 41 fathers (Table V). Sample sizes for item-level analyses ranged from 33 to 89 because some items were not applicable to parents of very young children (e.g., "I am comfortable talking with my child about his/her condition"). Additionally, items on the parent self-report Medication scale were only answered by parents of daughters with CAH (46,XX DSD) who were prescribed daily medication ($n = 33$).

Item-Level Analyses

Multitrait analyses with mothers' responses indicated that most items achieved corrected item-total correlations above the 0.40 threshold, and most items were more strongly correlated with their intended versus competing scales. Four items on the QOL-DSD-Proxy had corrected item-total correlations below .40, and one item was more strongly correlated with a competing than intended scale (Supplementary Data 1). For the QOL-DSD-Parent, eight items were below the .40 threshold, and eight items were more strongly correlated with competing versus their intended scales (Supplementary Data 2).

Item-Level Revisions

Based on these item-level analyses, subsequent revisions fell into three categories: (1) moving items to a different scale, (2) removing items owing to their psychometric properties, and (3) removing items owing to redundancy. In particular: (1) "My child's condition affects his/her activities" on the QOL-DSD-Proxy was moved from Social to Physical Functioning because of a higher loading; (2) items that did not fit on any scale and were rarely endorsed were removed (one DSD-Proxy item; four DSD-Parent items; e.g., "I worry about having another child with the same condition"), and (3) redundant items were removed (DSD-Parent: e.g., "I feel comfortable talking with my spouse about my child's condition" was removed and "...comfortable talking with family members about my child's condition" was retained).

The QOL-DSD-Proxy Social and Emotional Functioning scales were combined into a Socio-Emotional Functioning scale because of their high inter-correlations. Finally, four items on the QOL-DSD-Parent were retained on their original scales because the item content was more appropriate (e.g., "I felt stressed talking with my child before surgery" retained on the Surgery scale despite a high correlation with the Medications scale).

Given that the original QOL-DSD-Parent was 67 items, it was important to delete redundant questions to reduce respondent burden. In contrast, we retained some items that appeared to be central to having a child with a DSD from the qualitative data, despite their lower item-to-total correlations (e.g.,

"My child urinates differently than other children," "I feel stressed about health insurance and medical costs"). The final QOL-DSD-Proxy consists of 24 items, and the QOL-DSD-Parent consists of 55 items. Items (including removed items, denoted with a superscript) are listed in Supplementary Data 1 and 2.

Scale-Level Analyses

Scale-level analyses were computed on retained items (Table V). Most scales evidenced minimal to no floor effects. On the QOL-DSD-Proxy, negligible floor effects were found and ceiling effects ranged from 5.63% (Medical Concerns) to 32.88% (Gender Expression). On the QOL-DSD-Parent, floor effects (i.e., poor HRQoL) were found on the Surgery (11.54%) and Early Experiences (28.09%) scales, suggesting that nearly one-third of parents experienced significant distress during the diagnosis period. Ceiling effects were found on 7 out of 12 parent self-report scales; largest ceiling effects were found on the Social Functioning (55.06%) and Gender Expression (70.45%) scales. Thus, parents reported few negative effects on their *own* social functioning and few concerns about gender-related issues.

Reliability

Across both measures, internal consistency was adequate to good. Internal consistency on the QOL-DSD-Proxy was found for most scales, with Cronbach's alphas ranging from .53 (Physical Functioning) to .85 (Socio-Emotional Functioning). Internal consistency for the Physical Functioning scale improved slightly when analyses were restricted to children aged 4 and above ($\alpha = .59$) or to children who had had two or more surgeries ($\alpha = .59$). However, this coefficient was below the conventional cut-off of .70.

The QOL-DSD-Parent evidenced stronger internal consistency than the proxy measure, with alphas above .70 for 11 out of 12 scales (Table V). For children who had surgery or clinic visits within the past year, internal consistency for the Surgery and Doctor's Visits scales were better than for those who had these events >1 year ago (events within past year; α's = .74 and .80, respectively).

Test–retest reliability was available on the proxy measure for mothers of 21 randomly selected children (mean age = 4.14 years, $SD = 1.63$, mean genital atypicality = 44.89, $SD = 21.09$; 44.8% boys). ICCs were strong, ranging from .65 (Gender Expression) to .90 (Socio-Emotional Functioning; Table V). For the self-report measure, test–retest data were available for 25 randomly selected mothers (mean child age = 4.24 years, $SD = 1.62$; Mean genital atypicality = 41.12, $SD = 21.85$; 51.4% boys). ICCs for 13 out of 14 scales

were adequate to strong, ranging from .61 (Doctor's Visits) to .97 (Role Functioning; Table V). The Gender Expression scale had a low ICC of .11 in the overall sample; however, ICCs were more stable for boys (ICC = .43) and for those with greater atypicality in their genital appearance (ICC = .63).

Convergent Validity

Convergent validity for the QOL-DSD-Proxy was supported by correlations between this instrument with genital atypicality and PedsQL scores (Table VI). Although 32 parents consented to teacher participation, only 17 forms with corresponding parent measures were returned. A sample size of 34 would have been required to detect correlations stronger than 0.4, assuming 80% power (G*Power v3.1). Thus, we were underpowered to examine associations between the QOL-DSD and TRF.

Evidence of convergent validity was found for the QOL-DSD-Parent measure (Table VII). For a majority of scales, the hypothesized associations were found. For example, number of surgical procedures was negatively associated with the Earliest Experiences domain. The Decision-Making scale was moderately correlated with ERQ and DRS, and strongly correlated with the TiPS and all scales of the DCS. Notably, the Communication and Information and Disclosure scales were correlated with multiple scales of the HSCL, PSI, ERQ, IoF, DRS, DCS, and TiPS. Finally, convergent validity was documented for the QOL-DSD-Parent Medication, Surgery, Doctor's Visits, and Earliest Experiences scales.

Known-Group Differences

Known-group differences were tested for assigned gender (girls vs. boys), less versus greater genital atypicality, and for uncertainty versus no delay or change in gender assignment. As expected, greater genital atypicality (>50) was associated with significantly poorer Physical Functioning on the proxy measure, Welch statistic (due to heterogeneity in variances; 1, 69.22) = 11.46, $p < .001$. A change or delay in gender assignment was associated with worse Physical Functioning compared with those without a change or delay, $F(1,70) = 4.33$, $p < .05$.

For self-report, parent's perceptions of Gender Expression were significantly lower for parents of girls than parents of boys, $F(1,79) = 7.61$, $p < .01$, whereas parents of boys reported worse HRQoL on the Surgery scale, $F(1,69) = 4.24$, $p < .05$. The self-report measure also discriminated based on delay/change in gender assignment, which was associated with lower scores on the Earliest Experiences scale, Welch statistic (due to heterogeneity in variances; 1, 65.03) = 4.58, $p < .05$.

Discussion

There is a critical need to evaluate the impact of clinical procedures for managing DSD, assess long-term outcomes from the patients' and parents' perspectives, and provide standardized quality of life information to guide ongoing improvements in clinical care (Creighton, Michala, Mushtaq, & Yaron, 2013; Lantos, 2013; Lee et al., 2006; Roen & Pasterski, 2014). This is the first study to develop and evaluate DSD-specific HRQoL measures in young children (via a parent proxy) and their parents. Both measures demonstrated adequate to good psychometric properties, including internal consistency, test–retest reliability, convergent validity, and detection of known-group differences. These measures can be used to monitor health outcomes as clinical management changes, identify unmet clinical needs, and evaluate new interventions from a patient/family-centered perspective.

Experiences of Young Children With DSD and Their Parents

Our study provided a rare opportunity to characterize the experiences of children with DSD and their parents, and the impact of these conditions on daily functioning. Overall, greatest distress was reported for Early Experiences, Surgery, and Future Concerns. The magnitude of stress was highest when thinking about past or future events. When recalling the time surrounding diagnosis and subsequent surgical or medical interventions, one-third of parents reported the maximum distress that could be endorsed. For example, stress related to the diagnosis and uncertainty about the child's gender was consistently elevated in this sample, and reconciling conflicting information about their child's condition was an ongoing concern. Stressors also included feeling overwhelmed and irritable, disclosing the condition to others, and finding daycare providers and babysitters. Taken together, these findings underscore the need for emotional support for parents during the diagnostic period, in addition to ongoing parent–provider communication and shared decision-making.

Most parents reported that their children displayed minimal distress in relation to gender expression, and most parents (70%) reported no distress in this area. However, these issues may become more salient as children enter elementary school and transition into puberty. As evidenced in popular media and legislative efforts (e.g., creating gender-neutral bathrooms, zero-tolerance bullying policies), changing attitudes toward gender-variant expressions may be reflected in our findings (Janssen & Erickson-Schroth, 2013; Vance, Ehrensaft, & Rosenthal, 2014; Walch, Ngamake, Francisco, Stitt, & Shingler, 2012). It is also possible that parents are less concerned about atypical gender behavior in children at this age. More research is

Table VII. QOL-DSD-Parent Convergent Validity

Measure	Item/Scale	Decision-Making (n)	(r)	Role Functioning (n)	(r)	Gender Expression (n)	(r)	Social Functioning (n)	(r)	Emotional Functioning (n)	(r)	Future Concerns (n)	(r)	Healthcare Comm/Info (n)	(r)	Disclosure (n)	(r)	Medications (n)	(r)	Surgery (n)	(r)	Doctor's Visits (n)	(r)	Earliest Experiences (n)	(r)
Chart	Child Age (years)	89	−0.06	85	0.06	88	0.26*	89	0.14	89	0.13	89	−0.16	87	−0.03	88	−0.11	33	0.12	78	0.12	87	−0.04	89	−0.15
	Genital Atypicality	87	0.08	83	0.06	86	0.06	87	0.01	87	0.03	87	−0.06	85	0.06	86	−0.01	32	−0.10	**76**	**−0.01**	85	−0.05	**87**	**−0.09**
	N surgeries	76	0.15	72	−0.17	75	0.17	76	−0.19	76	−0.03	76	−0.05	74	0.10	75	−0.17	27	−0.23	**65**	**−0.06**	75	0.05	**76**	**−0.24***
HSCL	Somatization	88	0.02	84	−0.14	87	−0.01	88	−0.26*	88	**−0.39****	88	−0.13	86	−0.10	87	−0.11	32	0.07	77	−0.27*	86	−0.24*	88	0.03
	Obsessive Compulsive	88	0.01	84	−0.17	87	0.03	88	−0.25*	88	**−0.45****	88	−0.15	86	−0.15	87	−0.21	32	−0.14	77	−0.24*	86	−0.28*	88	0.00
	Interpersonal Sensitivity	88	−0.06	84	−0.03	87	−0.03	88	−0.25*	88	**−0.53****	88	−0.13	86	−0.22*	87	−0.20	32	0.04	77	−0.25*	86	−0.21*	88	0.01
	Depression	88	0.00	84	−0.14	87	0.08	88	−0.35**	88	**−0.60****	88	−0.12	86	−0.21	87	−0.25*	32	−0.06	77	−0.30**	86	−0.21	88	−0.02
	Anxiety	88	−0.11	84	−0.19	87	−0.04	88	−0.32**	88	**−0.45****	88	−0.14	86	−0.25*	87	−0.26*	32	−0.08	77	−0.36**	86	−0.31**	88	−0.01
	Total score	88	−0.04	84	−0.15	87	0.00	88	−0.33**	88	**−0.56****	88	−0.16	86	−0.22*	87	−0.24*	32	−0.05	77	−0.33**	86	−0.28**	88	0.00
PSI	Parenting Role Restriction	88	−0.14	84	−0.54**	87	−0.37**	88	−0.55**	88	−0.53**	88	−0.33**	86	−0.40**	87	−0.38**	33	−0.63**	77	−0.33**	86	−0.51**	88	−0.29**
	Parenting Competence	88	−0.14	84	−0.32**	87	−0.04	88	−0.32**	88	−0.30**	88	−0.20	86	−0.22*	87	−0.30**	33	−0.10	77	−0.09	86	−0.25*	88	−0.14
ERQ	Total Score	**89**	**−0.38****	85	−0.44**	88	−0.21*	89	−0.34**	89	−0.36**	89	−0.63**	87	−0.59**	88	−0.55**	33	−0.37*	78	−0.29*	87	−0.40**	89	−0.34**
IoF	General Negative Impact	**88**	**0.28****	84	0.75**	87	0.40**	88	0.75**	88	0.51**	88	0.38**	86	0.47**	87	0.41**	32	0.60**	77	0.38**	86	0.53**	88	0.29**
DRS	Decisional Regret Score	**89**	**−0.36****	85	−0.33**	88	−0.19	89	−0.35**	89	−0.07	89	−0.21*	87	−0.40**	88	−0.18	33	−0.04	78	−0.05	87	−0.23*	89	−0.10
DCS	Past: Effective Decision-Making	**60**	**−0.63****	58	−0.13	59	0.05	60	−0.17	60	−0.06	60	−0.32*	59	−0.49**	60	−0.20	26	0.08	55	0.12	58	−0.02	60	−0.02
	Past: Total Score	**60**	**−0.48****	58	0.23	59	0.34**	60	0.14	60	0.21	60	−0.03	59	−0.09	60	0.06	26	0.53**	55	0.36**	58	0.30*	60	0.26*
	Current: Effective Decision-Making	**28**	**−0.53****	26	−0.11	28	−0.06	28	−0.11	28	−0.11	28	−0.12	27	−0.54**	27	−0.30	6	−0.21	22	−0.35	28	−0.33	28	0.08
	Current: Total Score	**28**	**−0.05**	26	0.20	28	0.13	28	−0.15	28	−0.29	28	0.12	27	−0.13	27	0.01	6	−0.39	22	−0.17	28	−0.18	28	0.22
TiPS	Total Score	**88**	**0.51****	84	0.17	87	0.25*	88	0.12	88	0.17	88	0.26*	86	0.57**	87	0.22*	32	−0.14	77	0.09	86	0.15	88	0.20

Note. HSCL = Hopkins Symptom Checklist; PSI = Parenting Stress Index; ERQ = Experiences and Reactions Questionnaire; IoF = Impact on Family Scale; DRS = Decisional Regret Scale; DCS = Decisional Conflict Scale; TiPS = Trust in Physician Scale.
Hypothesized relationships between scales are shown in bold. r = Pearson's r.
*p < .05, **p < .01.

needed to understand the longitudinal trajectory of adaptation to these conditions. Additional HRQoL measures will be needed for school-age children and adolescents.

Psychometric Properties of the QOL-DSD-Proxy and QOL-DSD-Parent

Parent Proxy-Report

Overall, the QOL-DSD-Proxy demonstrated good psychometric properties; however, some scales had limited reliability or validity. First, internal consistency for Physical Functioning was below established cut-offs. This is likely owing to the heterogeneity of DSD syndromes and their effects on physical functioning. Reliability would be improved by removing items that apply to a more limited range of DSD; however, the resulting measure would then only apply to certain conditions (e.g., CAH). Thus, there is a trade-off between achieving conventional levels of internal consistency and adequate representation of these heterogeneous conditions.

Convergent validity for the Physical Functioning scale was not supported using the PedsQL. This is likely because items on the PedsQL are generic and not applicable to children with DSD (e.g., difficulty walking or climbing stairs). In contrast, our measure was specifically designed to be sensitive to the distinctive experiences of this population and, accordingly, physical functioning items focused on challenges with voiding. We did find that scores on the QOL-DSD Physical Functioning scale successfully discriminated between high versus low genital atypicality scores, as well as between those with versus without a change or delay in gender assignment.

Finally, although the study was underpowered to detect relationships with teacher reports, correlations >.40 emerged that would likely be statistically significant with a larger sample. It should be noted that the TRF Adaptive Functioning scale was marginally related to Socio-Emotional Functioning. Thus, our measure is likely able to detect both challenges and resilience.

Parent Self-Report

The scale structure, reliability, and validity of the QOL-DSD-Parent were supported. Scales with longer recall windows (e.g., Surgery, Doctor's Visits) asked parents to think about the last time their child had a surgical procedure/medical visit, and reliability was unsurprisingly below established cut-offs. Notably, reliability improved for parents reporting these events within the past month. We retained these scales because they are clinically important for children and families who have experienced *recent* surgeries or doctor's visits, but recommend administration of these scales only when these events have occurred *within the past month*.

Although most self-report scales evidenced strong stability, the Gender Expression scale evidenced poor test–retest reliability. First, it is possible that skewness violated test–retest assumptions, making the ICC invalid. However, this is unlikely because log-transformations of the scaled scores did not improve this coefficient. It is more likely that "gendered" play behaviors are unstable over time and are context-dependent (e.g., solitary vs. group play; Goble, Martin, Hanish, & Fabes, 2012; Maccoby, 2002). Research has indicated that young boys and girls display greater overlap and variability in gendered behavior than older children, and this is especially true for girls (Green, Bigler, & Catherwood, 2004; Sandberg & Meyer-Bahlburg, 1994; Sandberg, Meyer-Bahlburg, Ehrhardt, & Yager, 1993). The finding that test–retest reliability was higher for boys supports this interpretation.

Limitations and Future Directions

This is the first study to develop and validate HRQoL measures for children with DSD and their parents in a diagnostically diverse national sample. The following limitations should be acknowledged. First, our sample largely reflects prevalence rates of various DSD conditions. Accordingly, the predominant diagnoses were 46,XX CAH and 46,XY DSD with a proximal (i.e., severe) hypospadias phenotype; therefore, we may have failed to fully capture experiences relevant for less-prevalent DSD. Second, only 7% of children in our study were identified as having a DSD ascertained within the past year, so we were unable to examine HRQoL in the months immediately following diagnosis. Larger consortium studies will be needed to adequately represent rarer conditions and the impact of a recent diagnosis.

Research is needed to examine these DSD quality of life instruments, including: (1) responsiveness to change, (2) longitudinal outcomes, and (3) development of HRQoL measures for school-age children, adolescents, and adults with DSD. Qualitative interviews with children ages 6 and older are necessary to develop a self-report measure for school-age children. A measure for adolescents is particularly critical, because this period can involve uncertainties about body image, identity, capacity for forming intimate relationships, and increasing autonomy in healthcare (Sandberg et al., 2012). Together, these measures can be used to evaluate the effects of current clinical management strategies and new interventions, facilitate patient–provider communication, and promote shared decision-making.

References

Abidin, R. R. (1995). *Parenting Stress Index*. Odessa, FL: Psychological Assessment Resources, Inc.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Achermann, J., & Hughes, I. (2011). Disorders of sex development. In S. Melmed, K. Polonsky, P. Larsen, & H. Kronenberg (Eds.), *Williams textbook of endocrinology* (12th ed., pp. 868–934). Philadelphia, PA: W B Saunders Co.

Bell, B. A., Ferron, J. M., & Kromrey, J. D. (2008). Cluster size in multilevel models: The impact of sparse data structures on point and interval estimates in two-level models. *Proceedings of the Joint Statistical Meetings, Research Methods Section*. Alexandria, VA: American Statistical Association, pp. 1122–1129.

Berenbaum, S. A., Korman Bryk, K., Duck, S. C., & Resnick, S. M. (2004). Psychological adjustment in children and adults with congenital adrenal hyperplasia. *Journal of Pediatrics, 144*, 741–746.

Brehaut, J. C., O'Connor, A. M., Wood, T. J., Hack, T. F., Siminoff, L., Gordon, E., & Feldman-Stewart, D. (2003). Validation of a decision regret scale. *Medical Decision Making, 23*, 281–292.

Creighton, S. M., Michala, L., Mushtaq, I., & Yaron, M. (2013). Childhood surgery for ambiguous genitalia: Glimpses of practice changes or more of the same? *Psychology and Sexuality, 5*, 34–43. doi: 10.1080/19419899.2013.831214

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297–334.

Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral Science, 19*(1), 1–15. doi: 10.1002/bs.3830190102

Diamond, M., & Garland, J. (2014). Evidence regarding cosmetic and medically unnecessary surgery on infants. *Journal of Pediatric Urology, 10*, 2–6. doi: 10.1016/j.jpurol.2013.10.021

Feder, E. K. (2014). *Making Sense of Intersex: Changing Ethical Perspectives in Biomedicine*. Bloomington, IN: Indiana University Press.

Food and Drug Administration (2009). *Guidance for industry, patient-reported outcome measures: Use in medical product development to support labeling claims*. Retrieved from http://www.fda.gov/downloads/Drugs/GuidanceComplian ceRegulatoryInformation/Guidances/UCM193282.pdf. Retrieved March 5, 2014.

Goble, P., Martin, C., Hanish, L., & Fabes, R. (2012). Children's gender-typed activity choices across preschool social contexts. *Sex Roles, 67*, 435–451. doi: 10.1007/s11199-012-0176-9

Green, V., Bigler, R., & Catherwood, D. (2004). The variability and flexibility of gender-typed toy play: A close look at children's behavioral responses to counterstereotypic models. *Sex Roles, 51*, 371–386. doi: 10.1023/B:SERS.0000049227.05170.aa

Guyatt, G. H., Feeny, D. H., & Patrick, D. L. (1993). Measuring health-related quality of life. *Annals of Internal Medicine, 118*, 622–629.

Hays, R. D., Anderson, R., & Revicki, D. A. (1993). Psychometric considerations in evaluating health-related quality of life measures. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation, 2*, 441–449.

Hays, R. D., & Hayashi, T. (1990). Beyond internal consistency reliability: Rationale and user's guide for Multitrait Analysis Program on the microcomputer. *Behavior Research Methods, Instruments and Computers, 22*, 167–175.

Holmbeck, G. N., & Devine, K. A. (2009). Editorial: An author's checklist for measure development and validation manuscripts. *Journal of Pediatric Psychology, 34*, 691–696.

Hughes, I. A., Houk, C., Ahmed, S. F. & Lee, P. A., LWPES Consensus Group; ESPE Consensus Group. (2006). Consensus statement on management of intersex disorders. *Archives of Disease in Childhood, 91*, 554–563.

Janssen, A., & Erickson-Schroth, L. (2013). A new generation of gender: Learning patience from our gender nonconforming patients. *Journal of the American Academy of Child and Adolescent Psychiatry, 52*, 995–997. doi: 10.1016/j.jaac.2013.07.010

Karkazis, K. (2008). *Fixing sex: Intersex, medical authority, and lived experience*. Durham, NC: Duke University Press.

Kogan, B. A., Gardner, M., Alpern, A. N., Cohen, L. M., Grimley, M. B., Quittner, A. L., & Sandberg, D. E. (2012).

Challenges of disorders of sex development: Diverse perceptions across stakeholders. *Hormone Research in Paediatrics, 78*, 40–46.

Krupp, K., Fliegner, M., Brunner, F., Brucker, S., Rall, K., & Richter-Appelt, H. (2014). Quality of life and psychological distress in women with Mayer-Rokitansky-Küster-Hauser Syndrome and individuals with complete androgen insensitivity syndrome. *Open Journal of Medical Psychology, 3*, 212–221. doi: 10.4236/ojmp.2014.33023

Langberg, J. M., Epstein, J. N., Simon, J. O., Loren, R. E. A., Arnold, L. E., Hechtman, L., . . . Wigal, T. (2010). Parent agreement on ratings of children's attention deficit/hyperactivity disorder and broadband externalizing behaviors. *Journal of Emotional and Behavioral Disorders, 18*, 41–50. doi: 10.1177/1063426608330792

Lantos, J. D. (2013). The battle lines of sexual politics and medical morality. *Hastings Center Report, 43*, 3–4. doi: 10.1002/hast.147

Lee, P., Schober, J., Nordenstrom, A., Hoebeke, P., Houk, C., Looijenga, L., . . . Woodhouse, C. (2012). Review of recent outcome data of disorders of sex development (DSD): Emphasis on surgical and sexual outcomes. *Journal of Pediatric Urology, 8*, 611–615. doi: 10.1016/j.jpurol.2012.10.017

Lee, P. A., Houk, C. P., Ahmed, S. F., & Hughes, I. A.; Participants in the International Consensus Conference on Intersex. (2006). Consensus statement on management of intersex disorders. *Pediatrics, 118*, e488–e500.

Lux, A., Kropf, S., Kleinemeier, E., Jurgensen, M., & Thyen, U. (2009). Clinical evaluation study of the German network of disorders of sex development (DSD)/intersexuality: Study design, description of the study population, and data quality. *BMC Public Health, 9*, 110.

Maccoby, E. E. (2002). Gender and group process: A developmental perspective. *Current Directions in Psychological Science, 11*, 54–58. doi: 10.1111/1467-8721.00167

McHorney, C. A., Ware, J. E., & Raczek, A. E. (1993). The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical Care, 31*, 247–263.

Migeon, C. J., Wisniewski, A. B., Brown, T. R., Rock, J. A., Meyer-Bahlburg, H. F., Money, J., & Berkovitz, G. D. (2002). 46,XY intersex individuals: Phenotypic and etiologic classification, knowledge of condition, and satisfaction with knowledge in adulthood. *Pediatrics, 110*, e32.

Moseley, K. L., Clark, S. J., Gebremariam, A., Sternthal, M. J., & Kemper, A. R. (2006). Parents' trust in their child's physician: Using an adapted trust in physician scale. *Ambulatory Pediatrics, 6*, 58–61.

Muhr, T., & Freise, S. (2004). *User's manual for ATLAS.ti* (Version 5.0. 2). Berlin, Germany: Scientific Software Development.

National Institute of Diabetes and Digestive and Kidney Diseases. (2006, February). *Research progress report - strategic plan for pediatric urology*. Bethesda, MD: NIH Publication No. 06-5879.

Nunnally, J. C., & Berstein, I. H. (1994). *Psychometric theory*. New York, NY: McGraw-Hill.

O'Connor, A. M. (1995). Validation of a decisional conflict scale. *Medical Decision Making, 15*, 25–30.

Pasterski, V., Mastroyannopoulou, K., Wright, D., Zucker, K., & Hughes, I. (2014). Predictors of posttraumatic stress in parents of children diagnosed with a disorder of sex development. *Archives of Sexual Behavior, 43*, 369–375. doi: 10.1007/s10508-013-0196-8

Pasterski, V., Zucker, K. J., Hindmarsh, P. C., Hughes, I. A., Acerini, C., Spencer, D., . . . Hines, M. (2015). Increased cross-gender identification independent of gender role behavior in girls with congenital adrenal hyperplasia: Results from a standardized assessment of 4- to 11-year-old children. *Archives of Sexual Behavior, 44*, 1363–1375. doi: 10.1007/s10508-014-0385-0

Quigley, C. A., De Bellis, A., Marschke, K. B., el-Awady, M. K., Wilson, E. M., & French, F. S. (1995). Androgen receptor defects: Historical, clinical, and molecular perspectives. *Endocrine Reviews, 16*, 271–321.

Quittner, A. L., Davis, M. A., & Modi, A. C. (2003). Health-related quality of life in pediatric populations. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (pp. 696–709). New York, NY: Guilford Publications.

Roen, K., & Pasterski, V. (2014). Psychological research and intersex/DSD: Recent developments and future directions. *Psychology and Sexuality*, *5*, 102–116. doi: 10.1080/19419899.2013.831218

Rolston, A. M., Gardner, M., Vilain, E., & Sandberg, D. E. (2015). Parental reports of stigma associated with child's disorder of sex development. *International Journal of Endocrinology, 2015*, Article ID 980121, 15 pages. doi:10.1155/2015/980121

Sandberg, D., Gardner, M., & Cohen-Kettenis, P. (2012). Psychological aspects of the treatment of patients with disorders of sex development. *Seminars in Reproductive Medicine, 30*, 443–452. doi: 10.1055/s-0032-1324729

Sandberg, D., Gardner, M., & Rolston, A. (2010). *Experiences and Reactions Questionnaire*. Ann Arbor, MI: University of Michigan.

Sandberg, D. E., & Mazur, T. (2014). A noncategorical approach to the psychosocial care of persons with DSD and their families. In B. P. C. Kreukels, T. D. Steensma, & A. L. C. de Vries (Eds.), *Gender dysphoria and disorders of sex development* (pp. 93–114). New York, NY: Springer.

Sandberg, D. E., & Meyer-Bahlburg, H. F. (1994). Variability in middle childhood play behavior: Effects of gender, age, and family background. *Archives of Sexual Behavior, 23*, 645–663.

Sandberg, D. E., Meyer-Bahlburg, H. F., Ehrhardt, A. A., & Yager, T. J. (1993). The prevalence of gender-atypical behavior in elementary school children. *Journal of the American Academy of Child & Adolescent Psychiatry, 32*, 306–314.

Sanders, C., Carter, B., & Goodacre, L. (2011). Searching for harmony: Parents' narratives about their child's genital ambiguity and reconstructive genital surgeries in childhood. *Journal of Advanced Nursing, 67*, 2220–2230. doi: 10.1111/j.1365-2648.2011.05617.x

Sanders, C., Carter, B., & Goodacre, L. (2012). Parents need to protect: Influences, risks and tensions for parents of prepubertal children born with ambiguous genitalia. *Journal of Clinical Nursing, 21*, 3315–3323. doi: 10.1111/j.1365-2702.2012.04109.x

Schober, J., Nordenström, A., Hoebeke, P., Lee, P., Houk, C., Looijenga, L., ... Woodhouse, C. (2012). Disorders of sex development: Summaries of long-term outcome studies. *Journal of Pediatric Urology, 8*, 616–623. doi: http://dx.doi.org/10.1016/j.jpurol.2012.08.005

Schönbucher, V., Schweizer, K., Rustige, L., Schutzmann, K., Brunner, F., & Richter-Appelt, H. (2012). Sexual quality of life of individuals with 46,XY disorders of sex development. *The Journal of Sexual Medicine, 9*, 3154–3170. doi: 10.1111/j.1743-6109.2009.01639.x

Schroeder, J. F., Hood, M. M., & Hughes, H. M. (2010). Inter-parent agreement on the syndrome scales of the Child Behavior Checklist (CBCL): Correspondence and discrepancies. *Journal of Child and Family Studies, 19*, 646–653. doi: 10.1007/s10826-010-9352-0

Schwarz, N., & Sudman, S. (1996). *Answering questions: Methodology for determining cognitive and communicative processes in survey research*. San Francisco, CA: Jossey-Bass Publishers.

Spilker, B. (1996). *Quality of life and pharmacoeconomics in clinical trials* (2nd ed.). Philadelphia, PA: Lippincott-Raven Publishers.

Stein, R. E. K., & Riessman, C. K. (1980). The development of an impact-on-family scale: Preliminary findings. *Medical Care, 18*, 465–472.

Stout, S. A., Litvak, M., Robbins, N. M., & Sandberg, D. E. (2010). Congenital adrenal hyperplasia: Classification of studies employing psychological endpoints. *International Journal of Pediatric Endocrinology, 2010*, 11 pages. http://www.hindawi.com/journals/ijpe/2010/191520/. doi:10.1155/2010/191520

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. San Francisco, CA: Jossey-Bass.

van der Zwan, Y. G., Callens, N., van Kuppenveld, J., Kwak, K., Drop, S. L., Kortmann, B., ... ; Dutch Study Group on, D. S. D. (2013). Long-term outcomes in males with disorders of sex development. *The Journal of Urology, 190*, 1038–1042. doi: 10.1016/j.juro.2013.03.029

Vance, S. R., Ehrensaft, D., & Rosenthal, S. M. (2014). Psychological and medical care of gender nonconforming youth. *Pediatrics, 134*, 1184–1192. doi: 10.1542/peds.2014-0772

Varni, J. W., Seid, M., & Kurtin, P. S. (2001). Reliability and validity of the Pediatric Quality of Life Inventory Version 4.0 Generic Core Scales in healthy and patient populations. *Medical Care, 39*, 800–812.

Varni, J. W., Seid, M., & Rode, C. A. (1999). The PedsQL: Measurement model for the pediatric quality of life inventory. *Medical Care, 37*, 126–139.

Walch, S. E., Ngamake, S. T., Francisco, J., Stitt, R. L., & Shingler, K. A. (2012). The attitudes toward transgendered individuals scale: Psychometric properties. *Archives of Sexual Behavior, 41*, 1283–1291. doi: 10.1007/s10508-012-9995-6

Wisniewski, A. B., Migeon, C. J., Gearhart, J. P., Rock, J. A., Berkovitz, G. D., Plotnick, L. P., ... Money, J. (2001). Congenital micropenis: Long-term medical, surgical and psychosexual follow-up of individuals raised male or female. *Hormone Research, 56*, 3–11.

Wit, J. M., Ranke, M. B., & Kelnar, C. J. H. (2007). ESPE classification of paediatric endocrine diagnoses. *Hormone Research, 68*(Suppl. 2), 1–119.

Wolfe-Christensen, C., Fedele David, A., Mullins Larry, L., Lakshmanan, Y., & Wisniewski Amy, B. (2014). Differences in anxiety and depression between male and female caregivers of children with a disorder of sex development. *Journal of Pediatric Endocrinology and Metabolism, 27*, 617–621. doi: 10.1515/jpem-2014-0102.