Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Mind the gap: Informant discrepancies across domains of the Pediatric Quality of Life Index for intensive outpatient youth mental health care

Aline Thorkelsson
University of Alberta
Kalee Jack
University of Alberta
Kathryn Lambert
University of Calgary
Whitney Hindmarch
University of Calgary
Daniel Sadler
Recovery Alberta
Erin McCabe
emccabe@ualberta.ca

University of Alberta

Research Article

Keywords: pediatric, patient-reported outcome measures, youth mental health, informant discrepancies, child psychiatry, proxy-reported outcome measures

Posted Date: July 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10360948/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Purpose

To explore the relationships between youth and caregiver scores on the Pediatric Quality of Life Inventory (PedsQL) in an intensive outpatient youth mental health services setting.

Methods

This is a secondary analysis of routinely collected PedsQL data. Informant discrepancies between youth self-report and caregiver proxy-report were described. Differences across domains scores, and gender impacts on discrepancies were examined using Intra-Class Correlations and repeated measures ANOVA. The associations between age and discrepancies were examined using Spearman's rho. A hypothesis testing approach was used to assess the validity of interpreting discrepancies between scores.

Results

205 youth-caregiver dyads met eligibility criteria. The mean age was 14.8 +/- 1.7 and participants identified as boys (25.9%), girls (65.4%), and gender diverse (8.6%). ICC point estimates indicated poor agreement for all domains and global scores (0.378–0.487). Average absolute discrepancies varied from 12.0-17.3 across domains. There was a significant main effect of domain ($p = 0.018$) on discrepancies, but post-hoc tests were not significant. Age was significantly negatively associated with emotional and social domain discrepancies (ρ=-0.187,-0.123, respectively), and gender had no significant effect on discrepancies ($p = 0.867$). For construct validity, our hypotheses were supported by the data.

Conclusion

Youth and caregiver PedsQL scores show poor agreement. Informant discrepancies are highly variable across dyads, but do not show a bias towards youth or caregivers consistently under- or over-reporting quality of life. Age appears to influence discrepancies, but the impact of gender is unclear. The validity of interpreting informant discrepancies is supported; however, more research is needed to guide clinical interpretation.

Plain English Summary

1. Why is this study needed? Past studies found differences in youth and caregiver reports for standardized questionnaires that measure mental health symptoms and progress during treatment from their point of view. These differences are poorly understood for tools that measure general wellbeing in youth mental health settings like the Pediatric Quality of Life Inventory (PedsQL). Clinicians may use these scores to guide treatment, so an understanding of differences is important for patient care.

2. What is the key problem/issue/question this manuscript addresses? We investigated differences in youth and caregiver reports for the PedsQL and factors that may contribute to those differences.

3. What is the main point of your study? Differences do exist for youth- and caregiver-reported functioning using the PedsQL and these differences are not random.

4. What are your main results and what do they mean? There was a lot of variability in how large the differences were between youth and their caregivers for all categories of questions (emotional, physical, social, and school), and these differences are only partially explained by the youth's age. Differences for the emotional questions on the PedsQL strongly matched differences in another tool commonly used to measure anxiety and depression, which means the differences likely represent true differences in how youth and their caregivers perceive the youth's quality of life.

Introduction

Mental health disorders, including anxiety, depression, and behavioural challenges, are common among youth and place a significant burden on healthcare systems and families [1, 2]. In recent years, increasing attention has been given to incorporating the perspectives of youth into their own mental health care [3]. Patient-Reported Outcome Measures (PROMs) have emerged as a valuable tool in this context. These standardized, patient-centred questionnaires assess symptoms and functioning from the youth's perspective and are often completed by both youth and their caregivers. When used consistently, PROMs can support more accurate assessment, inform treatment decisions, and improve clinical outcomes [3–6]. Increasingly, policy agendas and clinical guidelines emphasize evaluating treatment impacts beyond symptom intensity by focusing on health-related quality of life (HRQOL), which captures how health conditions and treatments affect multiple aspects of a person's level of ability, daily functioning, and capacity to experience a fulfilling life.

A persistent challenge in the interpretation of PROM data is the low to moderate agreement between youth and caregiver reports, a phenomenon known as informant discrepancies [7]. In youth mental health care, these discrepancies are thought to follow predictable patterns that reflect differences in perceptions of internal experiences, insight, and communication [8, 9]. For example, it is thought that youth provide more accurate accounts of their subjective *internal* experiences, such as sadness, worry, or hopelessness, while caregivers are more accurate in reporting *external*, observable behaviours, such as behavioural problems [8, 10, 11]. In practice, this pattern is often interpreted as caregivers underestimating youth's internalizing symptoms, whereas youth-caregiver agreement is stronger for externalizing or more observable difficulties [7, 12]. Discrepancies also appear to widen during adolescence, which is when youth become more independent and are less likely to disclose emotional difficulties to caregivers [13]. Discrepancies can also be influenced by a range of factors, including the youth's age, family dynamics, and symptom visibility [13, 14].

Despite existing evidence, informant discrepancies are sometimes misinterpreted as random measurement error or attributed to the limitations of PROMs. This misinterpretation may obscure

meaningful differences in how symptoms or functional limitations are experienced or perceived by two individuals, both of whom report the same youth's mental health from their own distinct perspectives. It has been suggested that the discrepancy between a youth and their caregivers could be meaningfully interpreted as a third data point in youth mental health treatment [15]. If discrepancies are ignored during interpretations of multi-informant PROMs, clinical decisions informed by this data may be suboptimal, leading to a misalignment between treatment planning and the youth's needs. This risk is particularly high when clinician interpretation of scores relies exclusively on global scores, which can obscure important differences in caregiver and child perceptions within specific domains of HRQOL.

Informant discrepancies in symptom-based measures are well-documented, but fewer studies have examined these discrepancies in generic pediatric HRQOL instruments. Because HRQOL tools are generally multidimensional, they offer distinct insights into several related areas of functioning simultaneously, rather than rendering a single global summary of functioning. These investigations consistently find only moderate agreement, with emotional and school domains showing the weakest concordance [16–19]. One study of the PROMIS pediatric scales found moderate to low item-level agreement between youth- and caregiver-report and a pattern that aligns with higher discrepancies for internalizing (i.e., peer relationships, anger, anxiety and depression) versus externalizing symptoms (i.e., energy, fatigue, pain interference, asthma impact, upper extremity and mobility items) [20]. In a non-clinical sample responding to a general health questionnaire, youth also reported greater emotional difficulties than caregivers perceived, with behavioural-domain discrepancies being smaller [21].

The PedsQL is a widely used generic PROM in pediatric healthcare that is designed to assess HRQOL across four distinct domains physical, emotional, social, and school domains from both the youth and their caregiver's perspectives [22, 23]. Several studies have demonstrated the validity of the PedsQL for use in clinical youth mental health populations [10, 24, 25]. There is also research examining informant discrepancies within the PedsQL for general pediatric populations [17, 22, 24, 26]. One such study examined the impact of age on discrepancies, showing poor agreement for most psychosocial items only for older youth (ages 8–12) [17]. To our knowledge, no research has explored informant discrepancies on the PedsQL in youth mental health populations, the factors that may influence them, or how these discrepancies function across domain subscales versus global scores. A better understanding of informant discrepancies in PedsQL and other generic HRQOL instruments would not only enhance the validity of quality-of-life PROM score interpretation but also contribute to our understanding of how to integrate self- and proxy report HRQOL scores in research and service evaluation.

The overarching aim of this study was to explore the extent and nature of discrepancies between youth and caregiver scores on the PedsQL, moving beyond global metrics to examine its subscales. Specifically, the study had three objectives. First, we examined agreement and differences in both PedsQL global score and across HRQOL domains for youth and caregiver scores. Second, we tested whether discrepancies varied according to youth age and gender, and whether these patterns differed by

PedsQL domain. Finally, we used a construct validation approach to investigate the validity of interpreting youth-caregiver discrepancies in scores.

Methods

Design & Setting

This was a secondary analysis of routinely collected PROMs data from intensive, outpatient youth mental health services in Calgary, Canada. This study was approved by the University of Alberta Health Research Ethics Boards (Pro00147995).

Participants & Data collection

The sample included youth aged 8–18 years, along with their caregivers, who presented to a youth mental health centre in Calgary, Canada, between March 2023 and March 2025, for a severe mental health disorder requiring partial hospitalization or intensive outpatient care. Participants were treated in one of two services: 1) the Day Hospital program, a two-week intervention for youth transitioning from inpatient psychiatric care back into their communities, and 2) the Intensive Community Treatment Services (ICTS), a multidisciplinary program that provides short-term intensive support for youth with acute, escalating mental health concerns through a combination of group, family, and individual therapy, psychiatric care, and other psychosocial interventions. PROMs are administered to all youth aged 8 years or older and their caregivers at both intake and discharge. The measures were completed independently by both youth and their caregivers. Data included in this study were from intake only.

Demographic and clinical information, including age, gender, ethnicity, and clinical diagnosis were extracted from electronic health records. The clinical diagnoses represent the participants' principal administrative diagnosis. While the sample includes youth with neurodevelopmental conditions (e.g., Autism Spectrum Disorder, ADHD) or DSM-5 relational and adjustment codes, all participants met strict clinical thresholds for admission to the intensive outpatient mental health services due to severe, acute psychological distress and functional impairment. We performed a sensitivity analysis comparing the sample with mental health disorder diagnosis and neurodevelopmental disorder/relational or adjustment diagnosis and found no significant differences in absolute mean discrepancies across these groups (Supplementary File 1). Therefore, this sample thus represents an acute, transdiagnostic mental health treatment cohort. Item-level missingness did not occur as the electronic PROM system required all items to be complete at submission; however, cases missing an entire PedsQL domain were excluded from the analyses for that domain.

Measures

Two PROMs were administered across services at the centre: the acute version of the PedsQL and the Revised Children's Anxiety and Depression Scale (RCADS-25). These measures were selected to align with the most prevalent youth mental disorders, guided by the International Consortium for Health Outcomes Measurement's (ICHOM) Patient-Centred Outcome Set for Anxiety and Depression (including OCD and PTSD) [6, 27]. Choice of instruments was further informed by the PROMs available within the provincial electronic health record system and finalized through consultation with youth, caregivers, and clinicians.

Pediatric Quality of Life Inventory Generic Core Scales (PedsQL 4.0) [22]

The acute version of the PedsQL is a 23-item measure with a 7-day recall period that assesses four domains: physical, emotional, social, and school functioning. It has demonstrated construct validity in mental health populations [10, 23, 24]. Both youth self-report and caregiver-proxy forms are available, with youth forms divided into three age groups and caregiver forms into four age groups. Respondents rate the frequency of difficulties with various activities on a five-point scale. Items are reverse-scored and converted to a 0-100 scale, with higher scores reflecting better perceived HRQOL. For this study, we considered the emotional domain as capturing *internal* experiences (i.e., distress that is directed inward, manifesting as disturbances in mood), while social, school and physical domain captured external behaviours (i.e., more observable, objective domains of function).

Revised Children's Anxiety and Depression Scale- 25 (RCADS-25) [28, 29]

The RCADS-25 is a 25-item questionnaire for youth aged 8–18 years, comprising two internalizing symptoms subscales (anxiety and depression) as well as a global score. Items are rated on a 4-point frequency scale and yield age- and sex-normed T-scores [30]. Parallel self-report and caregiver-report versions are available [31]. The items ask about the frequency of symptoms on a 4-point scale [27]. The RCADS-25 has demonstrated strong reliability and validity in youth mental health populations, although its sensitivity to change has been less extensively studied [5].

Data Analysis

All analyses were conducted in JASP (version 0.19.3) [32]. Distributions and assumptions of normality for PedsQL data were assessed using a Shapiro-Wilk normalcy test and visual inspection of Q-Q plots.

Categorical variables (e.g., gender) were summarized as counts and percentages. Continuous variables (e.g., age) were summarized as means and standard deviations. Baseline PedsQL scores and RCADS-25 T-scores were summarized on the 0–100 scale for each domain and global score. Signed discrepancy scores for PedsQL global and domain scores were calculated as the difference between youth and caregiver proxy-report scores (youth - caregiver) for each domain. Positive PedsQL signed discrepancy values indicate higher self-reported functioning by the youth compared to their caregiver. Positive RCADS-25 signed discrepancy values corresponded to worse mental health symptom ratings by the

youth. Absolute discrepancies were calculated from signed discrepancies to show the magnitude of youth-caregiver differences.

Intraclass correlation coefficients (ICCs) were calculated using two-way mixed-effects models (ICC(3,1)) to assess absolute agreement between youth and caregiver reports. ICCs were calculated for global scores and separately for each domain to show if levels of agreement differed between internal (i.e., emotional domain) and external (i.e., physical, social and school) areas of functioning. A two-way repeated-measures ANOVA with rater (youth vs caregiver) and domain as within-subject factors was used to examine whether raw mean PedsQL scores differed for youth and caregivers and if there were differences across domains. Greenhouse-Geisser corrections were applied where appropriate, and Holm-corrected post-hoc tests were used to assess pairwise domain differences. Spearman's rank-order correlation coefficient was used to examine the association between age and signed discrepancy scores. A mixed repeated measures ANOVA was conducted to test if absolute value discrepancy scores differed according to domain (within-subjects) and gender (between-subjects). Greenhouse-Geisser corrections were applied if appropriate following sphericity testing, and Holm-corrected post hoc testing examined pairwise differences in discrepancy scores across domains and between genders.

Construct validity

We were interested in whether clinicians could meaningfully interpret the discrepancies between youth and their caregiver, that is, whether discrepancies represent true differences in how the youth and caregiver perceive the youth's quality of life in youth mental health. We therefore used a hypothesis testing approach to investigate whether absolute discrepancies in scores reflect true perceptual differences between a youth and their caregiver (which would be consistent across scales). We hypothesized that 1) there would be a moderate association (r = 0.3–0.5) between discrepancies across domains of the PedsQL, because one would expect the magnitude and direction of discrepancies in quality of life scores in one domain to relate to discrepancies in another quality of life domain; and 2) there would be a moderate-strong negative association between discrepancies on the PedsQL emotional domain and discrepancies in the RCADS-25 total score at intake (r = -0.5-0.7) as they capture similar constructs (i.e. emotional functioning and anxiety and depression symptoms). In previous work, we found a strong correlation between these two scores as a youth self-report [33]. Pearson correlations were used to test both hypotheses, examining associations between PedsQL discrepancy scores across domains and between PedsQL emotional domain discrepancy scores and RCADS-25 total scores. If these hypotheses were supported, it would be considered evidence for the validity of interpreting discrepancy scores as a third data point when using PedsQL score for informing clinical care in youth mental health.

Results

Participant Characteristics

A total of 600 electronic health records were screened, with 205 dyads met the eligibility requirements. Within this group, 3 dyads had missing data for the entire social domain. Participant characteristics are presented in Table 1. Over half of participants identified as girls (65.4%), 25.9% identified as boys, and 8.6% were gender diverse. The mean age at intake was 14.8 +/- 1.7 years. The most common ethnicity was Caucasian (69.6%), followed by Indigenous (5.9%) and Latino, Latina, or Latinx (3.7%). The most frequent diagnosis was anxiety (27.3%), followed by ADHD (20.5%), and depression (16.1%). Just over half of participants were attending ICTS (53.7%), with the remaining participants attending the Day Hospital (46.3%).

Shapiro-Wilk normalcy test p-values were < 0.001 for global mean absolute value discrepancy scores and for each domain at intake and discharge; however visual inspections of Q-Q plots suggested only minor deviations, so parametric tests were used.

Table 1
Participant characteristics.

Characteristic	
Age (years, mean +/- SD)	14.8 +/- 1.7
Gender (n = 162)	
Girls	106 (65.4%)
Boys	42 (25.9%)
Gender-diverse	14 (8.6%)
Ethnicity (n = 135)	
Caucasian	94 (69.6%)
Southeast Asian	6 (4.4%)
South Asian	3 (2.2%)
Middle Eastern	1 (0.7%)
Latino, Latina, Latinx	5 (3.7%)
Indigenous	8 (5.9%)
East Asian	2 (1.5%)
Black	1 (0.7%)
Mixed race	15 (11.1%)
Diagnosis (n = 161)	
ADHD	33 (20.5%)
Adjustment Disorder/difficulty coping	15 (9.3%)
Anxiety	44 (27.3%)
Autism Spectrum Disorder	3 (1.9%)
Depression	26 (16.1%)
Suicidal ideation	2 (1.24%)
Trauma- or Stressor-Related Disorder	4 (2.5%)
Caregiver or youth relational problem	14 (8.7%)
Borderline Personality Disorder	6 (3.7%)
Obsessive Compulsive Disorder	4 (2.5%)

Characteristic	
Other	10 (6.2%)
Service type (n = 205)	
Day Hospital	95 (46.3%)
Intensive Community Treatment Services	110 (53.7%)

Table 2 shows mean PedsQL and RCADS scores reported by youth and caregivers. global mean PedsQL scores were 47.7 (14.7) for youth and 47.6 (15.4) for caregivers. Figure 1 displays the youth and caregiver mean scores by domain.

Table 2
PedsQL and RCADS mean scores reported by youth and caregivers.

Scale	Youth report mean (SD)	Caregiver report mean (SD)
PedsQL		
Emotional	35.0 (16.8)	35.3 (14.5)
Physical	59.2 (20.9)	59.9 (23.4)
Social	56.56 (21.45)	53.2 (20.7)
School	33.66 (16.50)	35.0 (19.1)
Total	47.7 (14.7)	47.6 (15.4)
RCADS		
Anxiety	64.61 (20.98)	63.1 (28.4)
Depression	71.60 (21.40)	73.2 (31.3)
Total	71.68 (19.91)	72.9 (25.7)

Discrepancy Patterns

Table 3 shows means of the absolute values of the discrepancies and signed discrepancies. Mean absolute discrepancies were highest for the social domain, followed by the physical and school domains. Mean signed discrepancies were all small, with large standard deviations for global and all domain scores, with the social domain being the largest (3.4 +/- 22.4). Figure 2 shows the distribution of mean signed discrepancies by subscale.

Table 3
Mean absolute value discrepancies and mean signed discrepancies by domain at intake.

Domain	Mean absolute discrepancy	Mean signed discrepancy
Emotional (internal)	13.4 +/- 11.2	-0.4 +/- 17.5
Physical (external)	16.6 +/- 15.2	-0.6 +/- 22.6
Social (external)	17.3 +/- 14.5	3.4 +/- 22.4
School (external)	14.4 +/- 12.0	-1.2 +/- 18.7
Total	11.9 +/- 9.5	0.0 +/- 15.3

ICCs for youth- and caregiver-reported PedsQL scores by domain are shown in Table 4. Point estimates indicate poor agreement according to the criteria of Koo and Li (2016)[34] for all domains as well as global scores.

Table 4
ICC scores for mean youth- and caregiver-reported PedsQL scores by domain at intake.

Domain (type)	n	Point estimate	Lower 95% CI	Upper 95% CI
Emotional (internal)	205	0.378	0.255	0.490
Physical (external)	205	0.486	0.374	0.584
Social (external)	202	0.436	0.318	0.541
School (external)	205	0.459	0.344	0.560
Global Score	205	0.487	0.375	0.585

In terms of differences across domains, the ANOVA indicated no significant differences in the overall effect of the rater across domains: $F(1,200) = 4.499x10^{-5}$, $p = 0.995$, $\eta^2p = 2.249x10^{-7}$. There were overall differences across domain scores: $F(2.745,549.054) = 234.477$, $p = < 0.001$, $\eta^2p = 0.540$. Holm-corrected post-hoc testing showed statistically significant domain differences for all pairs ($p < 0.001$) except for emotional-school domain pair ($p = 0.391$). There was a small but statistically significant interaction effect between the rater and domain ($F(2.916,583.112) = 3.728$, $p = 0.012$). Holm-corrected post-hoc testing was not statistically significant ($p = 0.431–1.000$).

Age, Domain, and Gender Effects

Spearman's rho calculations revealed small but significant associations between age and global scores (ρ=-0.167, p = 0.017), as well as emotional (ρ=-0.187, p = 0.007), and social (ρ=-0.123, p = 0.007) domains, but not physical (ρ=-0.126, p = 0.072), or school (ρ=-0.034, p = 0.628). Table 5 contains the absolute and signed values for PedsQL global scores by gender. The mean global PedsQL absolute

discrepancies were 13.5 +/ -9.0 for girls, 10.9 +/- 9.3 for boys and 14.7 +/- 14.9 for gender diverse participants. Signed values were 0.8 +/- 15.7 for girls, 6.7 +/- 12.7 for boys, and − 7.2 +/- 20.0 for gender diverse participants. Mixed repeated measures ANOVA results for domain, gender, and the domain x gender interaction are shown in Table 6. Gender did not significantly predict discrepancy scores ($p = 0.867$). There was a significant main effect of domain ($F(3,468) = 3.375$, $p = 0.018$, $\eta^2p = 0.021$) on discrepancy scores, but Holm-corrected post-hoc tests did not identify statistically significant pairwise domain differences ($p = 0.057–1.000$). The domain x gender interaction was not statistically significant ($p = 0.957$).

Table 5
Mean absolute and signed discrepancies by gender.

Gender	Mean (SD) absolute discrepancy	Mean (SD) signed discrepancy
Girls (n = 106)	13.5 (9.0)	0.8 (15.7)
Boys (n = 46)	10.9 (9.3)	6.7 (12.7)
Gender diverse (n = 14)	14.7 (14.9)	-7.2 (20.0)

Table 6
Mixed repeated measures ANOVA for discrepancy absolute values looking at moderation by domain and gender.

Effect	Degrees of freedom	F	p	η^2p
Domain	3,468	3.375	0.018	0.021
Gender	2,156	0.143	0.867	0.002
Domain x gender	6,468	0.257	0.957	0.003

Construct validity

Pearson correlations between 1) PedsQL domain signed discrepancies and 2) PedsQL emotional domain mean discrepancies and RCADs-25 total scores are shown in Table 6. Discrepancy scores were moderately correlated for all domains, supporting our hypothesis that the PedsQL domain scores would correlate. There was a moderate-strong negative correlation for PedsQL emotional domain mean discrepancies and RCADS-25 total mean discrepancies ($r = -0.525$, $p < 0.001$), supporting our second hypothesis.

Table 6
Pearson correlations between PedsQL domain mean discrepancies and between PedsQL emotional domain mean discrepancies and RCADS-25 total scores to examine construct validity.

Pearson correlations	Pearson's r	p	Lower 95% CI	Higher 95% CI
Pearson correlations: PedsQL domains mean discrepancies				
Physical – emotional	0.369	< 0.001	0.244	0.482
Physical – social	0.483	< 0.001	0.370	0.582
Physical – school	0.374	< 0.001	0.250	0.486
Emotional – social	0.332	< 0.001	0.203	0.449
Emotional – school	0.386	< 0.001	0.263	0.497
Social – school	0.391	< 0.001	0.258	0.502
Pearson correlation: PedsQL emotional domain mean discrepancies and RCADS-25 total mean discrepancies				
PedsQL emotional – RCADS-25 total	-0.525	< 0.001	-0.617	-0.418

Discussion

This study investigated youth-caregiver informant discrepancies on the PedsQL within an intensive outpatient youth mental health setting. While the poor agreement between self- and proxy-report (ICCs between 0.378–0.487) aligns with broader PROMs research [16–19], our data revealed a near symmetrical balance of over- and under-reporting at the aggregate level. Youth and caregivers reported nearly identical group-level mean global quality of life scores (47.7 vs. 47.6, respectively), and mean domain scores, creating an illusion of alignment that could mask the variability across individual dyads. When evaluating the absolute magnitude of disagreement, large and clinically meaningful differences emerged across all PedsQL domains, ranging from 11.9 to 17.3 points. Given that established Minimal Important Difference thresholds for the PedsQL are around 5 points [35], the discrepancies observed within dyads are on average double (for global scores) to triple the threshold of clinical meaningfulness (for domain scores), peaking in the social domain. Notably, the emotional domain (classically expected to capture the greatest divergence because it is an internal experience) demonstrated the lowest mean absolute discrepancies.

These findings offer a notable departure from prior meta-analyses literature documenting a systematic caregiver underestimation of youth-reported internalizing symptoms (e.g., anxiety, depression), thought to be due to their low observability [7, 12]. This divergence suggests that informant discrepancies in youth mental health may operate differently for multi-dimensional quality-of-life outcomes than for symptom severity measures in youth mental health populations. Unlike symptom measures that track less observable internal states, the individual domains of PedsQL focus on perceived functioning and participation in daily life. As such, all domains, including emotional functioning, may be experienced and

interpreted as observable impacts on behavior. This suggests that individual appraisals and contextual factors (e.g., caregiver expectations, youth insight, family communication) may play a larger role in shaping discrepancies than domain-specific observability.

Taken together, these findings underscore the importance of interpreting both youth and caregiver HRQOL scores at the domain-specific level rather than relying on global scores to understand where a caregiver and their child differ in their perspectives on a child's HRQOL, and to avoid extending the traditional assumption that caregivers underestimate internalizing symptoms to the interpretation of emotional domain scores of the PedsQL.

Social-demographic factors associated with discrepancies revealed small but significant negative associations between age and the global scores, as well as the emotional and social domains. Negative correlations indicate that as youth age, they tend to experience worse emotional and social functioning than their caregivers perceive. Prior research suggests that informant discrepancies widen in adolescence [13], and specific to the PedsQL, poor agreement between child and caregiver on most psychosocial items was observed only for older children (ages 8–12) [17]. Prior. Our study extends these findings to mid-adolescence (mean age 14.7 years) and tentatively suggests that emotional and interpersonal functioning gaps may begin during pre-adolescence and continues to widen into the teenage years [15, 17]. Notably, age did not predict discrepancies in physical or school domains, reinforcing the notion that age-related perceptual gaps may be confined to specific domains of HRQOL, rather uniformly across the entire HRQOL instrument.

Gender was not a significant predictor of informant discrepancies, which is consistent with prior evidence that gender is not a primary driver of informant discrepancies [36–39]. The boys in our sample did appear to have a slight directionality in mean signed discrepancies (mean 6.7 (12.7)), meaning boys may rate their HRQOL higher compared to their caregivers, while in our sample of diverse gender participants, it appears that gender-diverse youth may experience lower overall HRQOL than perceived by their caregivers. This finding, alongside prior reports of caregivers underestimating the academic and social challenges faced by gender-diverse youth, suggests a need for further research into how gender-diverse youth and their caregivers perceive specific domains of HRQOL [40].

Our construct validation approach confirmed the validity of interpreting youth-caregiver discrepancies in scores as a source of information for clinical decision-making. As hypothesized, moderate associations between discrepancies across domains were found. Evidence also supported our hypothesis that discrepancies in the PedsQL emotional domain would have a moderate-strong association with discrepancies in RCADS-25 total scores. This finding means that although there were large variations in discrepancies across youth-caregiver dyads, discrepancies are similar across two scales capturing similar constructs (emotional functioning and anxiety/depression). Further, these discrepancies are consistent *within* the dyads across domain scores on the PedsQL. This study therefore supports the idea that informant discrepancies in generic HRQOL instruments are meaningful and should be attended to

as a third data point (i.e., youth, caregiver and discrepancy in scores) when PROMs are used as part of routine care [41].

When using multidimensional HRQOL measures like the PedsQL with multiple informants (self and proxy-report) in routine care, interpreting domain-level discrepancies may offer distinct clinical utility beyond merely documenting differing perspectives. The magnitude and direction of the discrepancy could be interpreted as an indicator of the caregiver-child alignment on specific HRQOL domains. For instance, a high absolute discrepancy confined to the social domain points to a localized breakdown in caregiver awareness or youth disclosure regarding peer interactions, whereas an isolated emotional domain discrepancy may reflect divergent appraisals of the child's emotional functioning. In addition, treating these domain-level discrepancies as a discrete metrics could allow researchers and health system administrators to evaluate whether an intervention or program successfully brings a family into perceptual alignment offering a more nuanced understanding of where programs are effective than global quality-of-life scores. Future studies should pair PedsQL discrepancy scores with concurrent measures of related constructs to understand their validity. For example, the validity of social domain discrepancies could be investigated with a combination of observational coding of parent-child communication, and validated child self-disclosure scales.

Limitations and Future Directions

The limitations of this study include small subgroup sizes, particularly among gender-diverse youth, and the subsequent reduced power for interaction effects, which limits the generalizability of sociodemographic-related findings. The observational design restricts causal inference, and analyses rely on routinely collected data, which may vary in the standardization of the PROMs administration. Use of routinely collected data also limited the number of variables available to examine as predictors of discrepancies. Finally, the sample was drawn from a single Canadian mental health centre, which may limit external validity.

Future research should focus on making discrepancies clinically interpretable for clinicians. For example, understanding the threshold of discrepancies at which youth-caregiver dyads would benefit from targeted psychoeducation or family-based interventions, investigating whether interventions explicitly targeting caregiver-youth relationships and communication can reduce discrepancies, and whether reductions in discrepancies predict positive long-term clinical outcomes. Longitudinal studies of change in informant discrepancies as children and youth age, and as they go through mental health treatment would also provide valuable information for interpreting discrepancies. Larger and more diverse samples are needed to clarify the roles of gender, age, and other contextual factors such as youth and caregiver level of insight, family dynamics, and caregiver expectations. Qualitative approaches could provide deeper insights into the mechanisms underlying diverging perceptions of HRQOL.

Conclusion

There is poor agreement between youth and their caregivers on their PedsQL scores, with high variation across youth-caregiver dyads. There were significant differences in the magnitude of discrepancies across HRQOL domains, which was influenced by age in the social and emotional domains, pointing to the need to interpret at the domain level rather than global HRQOL scores. Informant discrepancies are consistent within dyads, providing evidence of the validity of interpreting discrepancies as a data point in clinical care.

Declarations

Competing Interests:

The authors have no relevant financial or non-financial interests to disclose.

Ethics approval:

This was a secondary analysis of routinely collected PROMs data from child and adolescent mental health services in Calgary, Canada. This study was approved by the University of Alberta Health Research Ethics Boards (Pro00147995).

Consent to participate:

Not applicable.

Consent to publish:

Not applicable.

Statement on patient, public, or community engagement

Not applicable.

Funding:

This work was supported by the Stollery Children's Hospital Foundation through the Women and Children's Health Research Institute.

Author Contribution

EM, WH, AT, KL contributed to the study conception and design. Data cleaning and analyses were performed by AT, KJ and DS. The first draft was written by AT and KJ, and refined by EM, KL, WH, and DS. All authors read and approved the final manuscript.

Acknowledgement

This research was supported by a WCHRI summer studentship which has been funded by the Stollery Children's Hospital Foundation through the Women and Children's Health Research Institute. We wish to thank the administrative support, clinical staff and managers at The Summit for their support of this project.

Data Availability

The data for this study is available from the corresponding author with an approved data disclosure agreement from Recovery Alberta.

References

1. Ravens-Sieberer, U., Kaman, A., Erhart, M., Otto, C., Devine, J., Löffler, C., & Hölling, H. (2023). Quality of life and mental health in children and adolescents during the first year of the COVID-19 pandemic: results of a two-wave nationwide population-based study. *European Child and Adolescent Psychiatry*, *32*(4). https://doi.org/10.1007/s00787-021-01889-1
2. Mental Health Research Canada (n.d.). https://www.mhrc.ca/.
3. Gibbons, C., Porter, I., Gonçalves-Bradley, D. C., Stoilov, S., Ricci-Cabello, I., Tsangaris, E., & Valderas, J. M. (2021). Routine provision of feedback from patient-reported outcome measurements to healthcare providers and patients in clinical practice. *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.CD011589.pub2
4. Rognstad, K., Engell, T., Fjermestad, K., Wentzel-Larsen, T., & Kjøbli, J. (2025). Process and Implementation Elements of Measurement Feedback Systems: A Systematic Review. *Administration and Policy in Mental Health and Mental Health Services Research*, *52*(1). https://doi.org/10.1007/s10488-023-01325-3
5. Krause, K. R., Chung, S., Adewuya, A. O., Albano, A. M., Babins-Wagner, R., Birkinshaw, L., & Wolpert, M. (2021). International consensus on a standard set of outcome measures for child and youth anxiety, depression, obsessive-compulsive disorder, and post-traumatic stress disorder. The Lancet Psychiatry. https://doi.org/10.1016/S2215-0366(20)30356-4
6. Thapa Bajgain, K., Amarbayan, M., Wittevrongel, K., McCabe, E., Naqvi, S. F., Tang, K., & Santana, M. (2023). Patient-reported outcome measures used to improve youth mental health services: a systematic review. *Journal of Patient-Reported Outcomes*. https://doi.org/10.1186/s41687-023-00556-0
7. De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, *141*(4). https://doi.org/10.1037/a0038498

8. Achenbach, T. M. (2006). As others see Us clinical and research implications of cross-informant correlations for psychopathology. *Current Directions in Psychological Science*, *15*(2). https://doi.org/10.1111/j.0963-7214.2006.00414.x
9. Sourander, A., Helstelä, L., & Helenius, H. (1999). Parent-adolescent agreement on emotional and behavioral problems. *Social Psychiatry and Psychiatric Epidemiology*, *34*(12), 657–663. https://doi.org/10.1007/s001270050189
10. McCabe, E., Hindmarch, W., Bajgain, B., Jacob, J., Arnold, P. D., Ortega, I., & Zwicker, J. D. (2025). Validity evidence for the use of the Pediatric Quality of Life Inventory, the Revised Children's Anxiety and Depression Scale-25, and the Columbia-Suicide Severity Rating Scale in measurement-based care in intensive outpatient child and adolescent mental health care. *Quality of Life Research*, *34*(10). https://doi.org/10.1007/s11136-025-04037-3
11. Ebesutani, C., Chorpita, B. F., Higa-Mcmillan, C. K., Nakamura, B. J., Regan, J., & Lynch, R. E. (2011). A psychometric analysis of the revised child anxiety and depression scales-parent version in a school sample. *Journal of Abnormal Child Psychology*, *39*(2). https://doi.org/10.1007/s10802-010-9460-8
12. Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/Adolescent Behavioral and Emotional Problems: Implications of Cross-Informant Correlations for Situational Specificity. *Psychological Bulletin*. https://doi.org/10.1037/0033-2909.101.2.213
13. de Los Reyes, A., Goodman, K. L., Kliewer, W., & Reid-Quiñones, K. (2010). The Longitudinal Consistency of Mother-Child Reporting Discrepancies of Parental Monitoring and Their Ability to Predict Child Delinquent Behaviors Two Years Later. *Journal of Youth and Adolescence*, *39*(12). https://doi.org/10.1007/s10964-009-9496-7
14. Stuart, J., & Jose, P. E. (2012). The influence of discrepancies between adolescent and parent ratings of family dynamics on the well-being of adolescents. *Journal of Family Psychology*, *26*(6). https://doi.org/10.1037/a0030056
15. De Los Reyes, A., & Epkins, C. C. (2023). Introduction to the Special Issue. *Journal of Clinical Child and Adolescent Psychology*. https://doi.org/10.1080/15374416.2022.2158843. A Dozen Years of Demonstrating That Informant Discrepancies are More Than Measurement Error: Toward Guidelines for Integrating Data from Multi-Informant Assessments of Youth Mental Health.
16. Jafari, P., Bagheri, Z., Hashemi, S. Z., & Shalileh, K. (2013). Assessing whether parents and children perceive the meaning of the items in the PedsQLTM 4.0 quality of life instrument consistently: a differential item functioning analysis. *Global journal of health science*, *5*(5). https://doi.org/10.5539/gjhs.v5n5p80
17. Khanna, D., Khadka, J., Mpundu-Kaambwa, C., & Ratcliffe, J. (2023). Child–Parent Agreement in the Assessment of Health-Related Quality of Life Using the CHU9D and the PedsQLTM. *Applied Health Economics and Health Policy*, *21*(6), 937–947. https://doi.org/10.1007/s40258-023-00831-7
18. van de Looij-Jansen, P. M., Jansen, W., de Wilde, E. J., Donker, M. C. H., & Verhulst, F. C. (2011). Discrepancies between parent-child reports of internalizing problems among preadolescent

children: Relationships with gender, ethnic background, and future internalizing problems. *Journal of Early Adolescence*, *31*(3). https://doi.org/10.1177/0272431610366243

19. Stokes, M. A., Kornienko, L., Scheeren, A. M., Koot, H. M., & Begeer, S. (2017). A comparison of children and adolescent's self-report and parental report of the PedsQL among those with and without autism spectrum disorder. *Quality of Life Research*, *26*(3). https://doi.org/10.1007/s11136-016-1490-4
20. Varni, J. W., Thissen, D., Stucky, B. D., Liu, Y., Magnus, B., He, J., & DeWalt, D. A. (2015). Item-level informant discrepancies between children and their parents on the PROMIS® pediatric scales. *Quality of Life Research*, *24*(8). https://doi.org/10.1007/s11136-014-0914-2
21. Caqueo-Urízar, A., Urzúa, A., Villalonga-Olives, E., Atencio-Quevedo, D., Irarrázaval, M., Flores, J., & Ramírez, C. (2022). Children's Mental Health: Discrepancy between Child Self-Reporting and Parental Reporting. *Behavioral Sciences*, *12*(10). https://doi.org/10.3390/bs12100401
22. Varni, J. W., Seid, M., & Kurtin, P. S. (2001). PedsQL™ 4.0: Reliability and Validity of the Pediatric Quality of Life Inventory™ Version 4.0 Generic Core Scales in Healthy and Patient Populations. *Medical Care*, *39*(8). https://doi.org/10.1097/00005650-200108000-00006
23. O'Loughlin, R., Jones, R., Chen, G., Mulhern, B., Hiscock, H., Devlin, N., & Dalziel, K. (2024). Comparing the Psychometric Performance of Generic Paediatric Health-Related Quality of Life Instruments in Children and Adolescents with ADHD, Anxiety and/or Depression. *Pharmacoeconomics*, *42*(Suppl 1). https://doi.org/10.1007/s40273-024-01354-2
24. Bastiaansen, D., Koot, H. M., Bongers, I. L., Varni, J. W., & Verhulst, F. C. (2004). Measuring quality of life in children referred for psychiatric problems: Psychometric properties of the PedsQL™ 4.0 generic core scales. *Quality of Life Research*, *13*(2). https://doi.org/10.1023/B:QURE.0000018483.01526.ab
25. De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, *141*(4). https://doi.org/10.1037/a0038498
26. Coombes, L. H., Wiseman, T., Lucas, G., Sangha, A., & Murtagh, F. E. M. (2016). Health-related quality-of-life outcome measures in paediatric palliative care: A systematic review of psychometric properties and feasibility of use. *Palliative Medicine*. https://doi.org/10.1177/0269216316649155
27. ICHOM (n.d.). Anxiety, Depression, OCD & PTSD in Children & Young People. https://www.ichom.org/patient-centered-outcome-measure/anxiety-depression-ocd-and-ptsd-in-children-and-young-people/.
28. Chorpita, B. F., Yim, L., Moffitt, C., Umemoto, L. A., & Francis, S. E. (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: A revised child anxiety and depression scale. *Behaviour Research and Therapy*, *38*(8). https://doi.org/10.1016/S0005-7967(99)00130-8
29. Muris, P., Meesters, C., & Schouten, E. (2002). A brief questionnaire of DSM-IV-defined anxiety and depression symptoms among children. *Clinical Psychology and Psychotherapy*, *9*(6). https://doi.org/10.1002/cpp.347

30. Ebesutani, C., Reise, S. P., Chorpita, B. F., Ale, C., Regan, J., Young, J., & Weisz, J. R. (2012). The Revised Child Anxiety and Depression Scale-Short Version: Scale reduction via exploratory bifactor modeling of the broad anxiety factor. *Psychological Assessment*, *24*(4). https://doi.org/10.1037/a0027283
31. Posner, K., Brown, G. K., Stanley, B., Brent, D. A., Yershova, K. V., Oquendo, M. A., & Mann, J. J. (2011). The Columbia-suicide severity rating scale: Initial validity and internal consistency findings from three multisite studies with adolescents and adults. *American Journal of Psychiatry*, *168*(12). https://doi.org/10.1176/appi.ajp.2011.10111704
32. JASP Team (2026). JASP (version 0.19.3) [Computer software]. https://jasp-stats.org/
33. McCabe, E., Hindmarch, W., Bajgain, B., Jacob, J., Arnold, P. D., Ortega, I., & Zwicker, J. D. (2025). Validity evidence for the use of the Pediatric Quality of Life Inventory, the Revised Children's Anxiety and Depression Scale-25, and the Columbia-Suicide Severity Rating Scale in measurement-based care in intensive outpatient child and adolescent mental health care. *Quality of Life Research*, *34*(10). https://doi.org/10.1007/s11136-025-04037-3
34. Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2). https://doi.org/10.1016/j.jcm.2016.02.012
35. Varni, J. W., Burwinkle, T. M., Seid, M., & Skarr, D. (2003). The PedsQL^{TM}* 4.0 as a pediatric population health measure: Feasibility, reliability, and validity. *Ambulatory Pediatrics*, *3*(6). https://doi.org/10.1367/1539-4409(2003)003%3C0329:TPAAPP%3E2.0.CO;2
36. van der Ende, J., Verhulst, F. C., & Tiemeier, H. (2012). Agreement of informants on emotional and behavioral problems from childhood to adulthood. *Psychological Assessment*, *24*(2), 293–300. https://doi.org/10.1037/a0025500
37. De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, *141*(4), 858–900. https://doi.org/10.1037/a0038498
38. Booth, C., Moreno-Agostino, D., & Fitzsimons, E. (2023). Parent-adolescent informant discrepancy on the Strengths and Difficulties Questionnaire in the UK Millennium Cohort Study. *Child and Adolescent Psychiatry and Mental Health*, *17*(1), 57. https://doi.org/10.1186/s13034-023-00605-y
39. Rognli, E. W., & Fjermestad, K. W. (2023). Informant Discrepancy in Report of Parent-adolescent Conflict as a Predictor of Hopelessness among Depressed Adolescents: A Replication Study. *Journal of Child and Family Studies*, *32*(8), 2522–2531. https://doi.org/10.1007/s10826-023-02610-3
40. Bhattacharya, N., Budge, S. L., Pantalone, D. W., & Katz-Wise, S. L. (2020). Conceptualizing relationships among transgender and gender diverse youth and their caregivers. *Journal of Family Psychology*. https://doi.org/10.1037/fam0000815
41. De Los Reyes, A., Talbott, E., Power, T. J., Michel, J. J., Cook, C. R., Racz, S. J., & Fitzpatrick, O. (2022). The Needs-to-Goals Gap: How informant discrepancies in youth mental health assessments impact

service delivery. *Clinical Psychology Review*. https://doi.org/10.1016/j.cpr.2021.102114

Figures

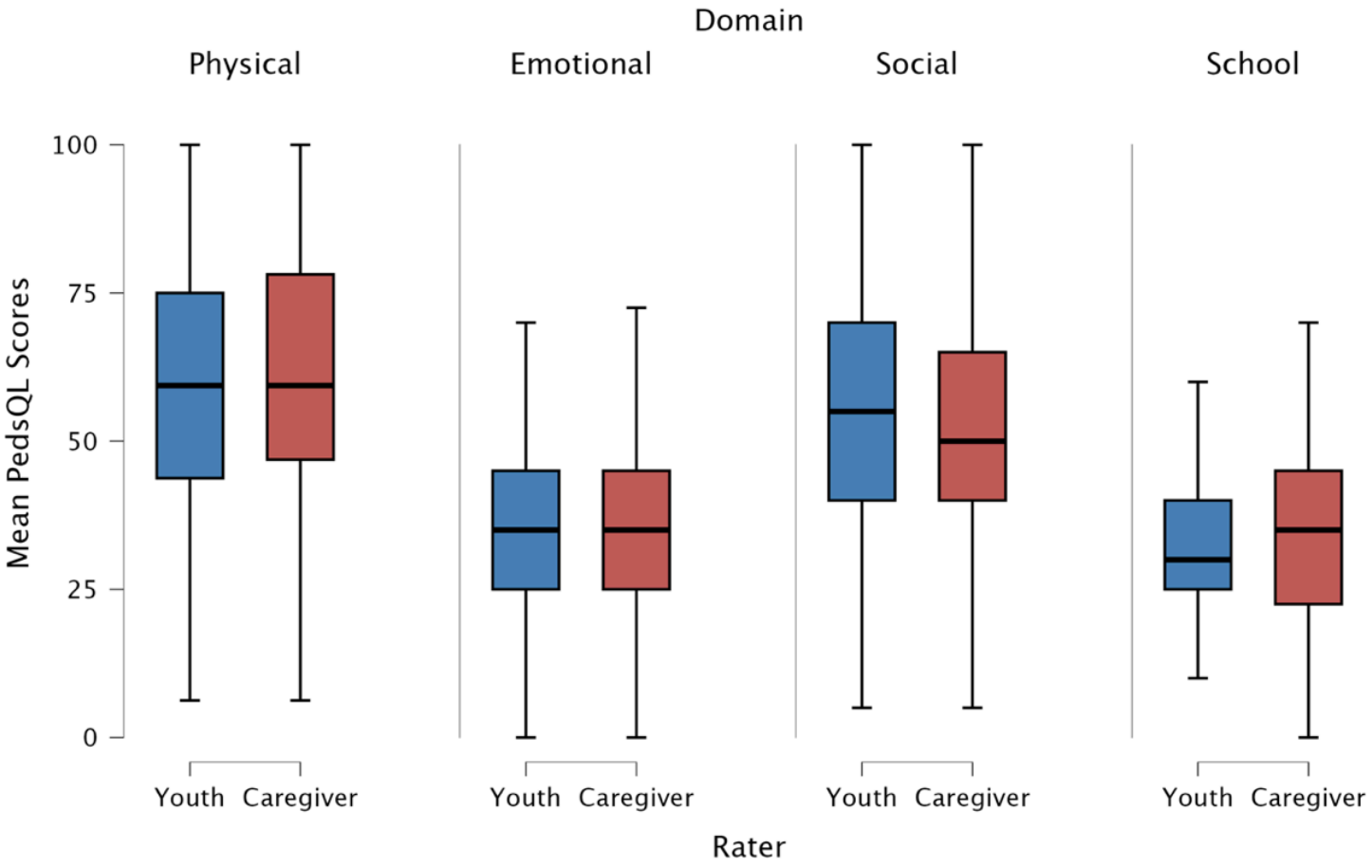


Figure 1

Youth and caregiver mean PedsQL scores by domain.

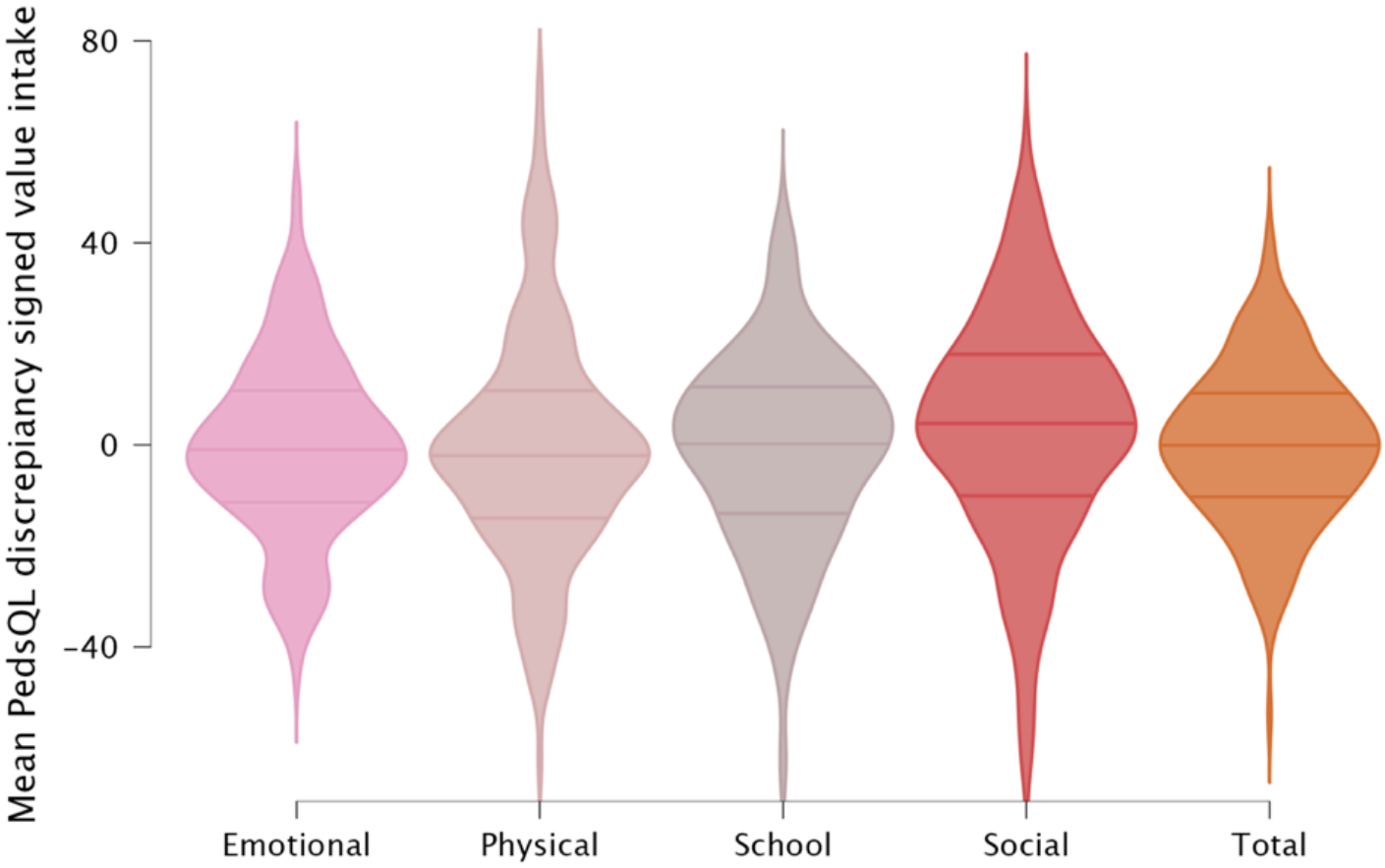


Figure 2

Violin plot of mean PedsQL signed discrepancy values at intake by domain and global values. Horizontal lines within the violin plot represent median, Q1, and Q3.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFile1.docx