Community Dent Oral Epidemiol 2007; 35 (Suppl. 1): 32–40
All rights reserved

Community Dentistry and Oral Epidemiology

Concordance between caregiver and child reports of children's oral health-related quality of life

Maureen Wilson-Genderson[1],
Hillary L. Broder[2] **and Ceib Phillips**[3]

[1]Department of Community Health, University of Medicine and Dentistry NJ, New Jersey Dental School, Newark, NJ, USA, [2]Department of Community Health, University of Medicine and Dentistry NJ, New Jersey Dental School, Newark, NJ, USA, [3]Department of Orthodontics, University of North Carolina, Chapel Hill, NC, USA

Wilson-Genderson M, Broder, HL, Phillips C. Concordance between caregiver and child reports of child's oral health-related quality of life. Community Dent Oral Epidemiol 2007; 35 (Suppl. 1): 32–40. © 2007 The Authors. Journal compilation © 2007 Blackwell Munksgaard

Abstract – *Objective:* This study sought to assess child-caregiver concordance regarding children's oral health related quality of life (OHRQoL) using the Child Oral Health Impact Profile (COHIP). *Methods:* The sample comprised treatment-seeking children aged 8–15 with pediatric (n = 141), orthodontic (n = 135), and craniofacial (n = 100) needs and their caregivers. Children and their caregivers were queried concerning the child's *Oral health, Functional Well-being, Social/Emotional Well-being, School environment and Self-image*. These combined subscales yielded an overall OHRQoL rating. The dyads were distributed at recruitment locations as follows: Montreal (50 pediatric, 13 orthodontic, 15 craniofacial), UMDNJ (45 pediatric, 15 orthodontic, 0 craniofacial), and NYU (46 pediatric, 107 orthodontic, 85 craniofacial). Concordance was assessed with Spearman and intraclass correlations and Kruskal–Wallis testing of categories of agreement. *Results:* Low to modest rates of agreement between child and caregiver were found for the sample overall. Rates of concordance between child and caregiver varied between clinical groups–craniofacial patients were more likely to rate OHRQoL higher than they were to agree with their caregivers' ratings. In contrast, pediatric and orthodontic patients were more likely either agree with or rate their OHRQoL lower than their caregivers' ratings. *Conclusion:* These findings of child-caregiver concordance using the COHIP supported previous work suggesting the usefulness of obtaining both child and caregiver reports of the child's QoL.

Key words: Child Oral Health Impact Profile (COHIP); child-caregiver concordance; clinical decision-making; craniofacial; oral health-related quality of life

Hillary Broder, Department of Community Health, University of Medicine and dentistry NJ, New Jersey Dental School, Newark, NJ 07101
Tel: +973 972 3612
Fax: +973 972 0363
e-mail: broder@umdnj.edu

All authors declare no conflicts of interest

The assessment of oral health-related quality of life (OHRQoL) ideally attempts to encompass 'how well or poorly life works at a particular time' (1) as a result of oral-facial health. It requires a rating of an individual's subjective experience regarding well-being or disease. An individual considering treatment for an oral-facial condition is often queried not only about the current experience of but also post-treatment expectations for OHRQoL. When the individual who potentially needs treatment is a child, an obvious concern is: who should provide the opinion on the child's quality of life and treatment expectations?

The child's opinion, it may be argued, is the most valuable opinion. However, a child may be too young or too ill to give an impression of his or her well-being. If the child is able to provide a self-report, the information may be subject to a few qualifications. The dominance of short-term memory, strong influence of recent incidents, absence of a fully developed long-term perspective, language problems (interview) and reading ability (written questionnaire) may impact the reliability and validity of the results or responses (2, 3).

For all these reasons, the usefulness of proxy reports has been investigated. Indeed, it is 'standard practice' to examine how well the proxy rating mirrors the child's rating when assessing a new QoL instrument (4). Although it may seem that a caregiver should adequately estimate the well-being of his or her child, there is a good deal of evidence indicating that caregivers generally have low to modest agreement with the child's rating (5, 6). Caregivers may over or under-estimate the importance of certain things like facial appearance, time away from school as well as symptomology and likely have biases and expectations that may influence the QoL rating (4). Further, caregivers do not observe their school-aged children throughout the entire day (e.g., school interactions, tooth brushing). They have been found to report higher QoL (7) as well as lower QoL (8) than the rated child.

Proxy reports have also been considered valuable because beyond being a possible substitute rating of QoL, they may enhance the understanding by providing 'a more comprehensive picture of the child across settings' (5). Thus, even if the rates of agreement among child-proxy reports are modest, such caregiver assessments could provide important additional information to guide treatment decisions. Teacher reports may also represent another proxy for the child (9).

Whether the proxy rating is to be used instead of or in addition to the child's own report, it is useful to consider how characteristics other than the child's health may be associated with the degree of agreement. The seriousness of the condition of the child has been examined as a factor that may be associated with whether the caregiver and child agree and in particular whether the parents of ill children rate their well-being lower than parents of well children. Some evidence suggests that parental QoL assessments of children with chronic health conditions or extreme conditions are more discrepant (when compared to the children's assessments) than when the child is healthy (10) and that the caregiver tends to report more negative impact than the child does him or herself (8, 11).

In addition to the physical health of the child, sex, age or race/ethnicity might be expected to affect the degree of agreement between the child and caregiver's ratings of the child's QoL. It has been suggested that agreement might improve, for example, as children moved into adolescence. Yet, Eiser and Morse (4) reviewed the findings regarding age and sex and found no consistent pattern of influence. This overall finding has led them and others (11) to suggest that it is important to examine how demographic and illness characteristics affect proxy ratings for specific disease states and for specific measures of HRQoL.

Considerable effort has been invested into developing instruments designed to measure oral health-related quality of life (OHRQoL) (12, 13) with recent effort aimed at developing a measure of oral health-related quality of life that is appropriate for use with children (14, 15).

In order to develop such a measure questions that must be addressed include: who is going to be the one to give an assessment of how well or poorly life is working for the child and how will those ratings be viewed? To date, scant data are available when comparing children across oral health conditions. This paper presents an evaluation of the inter-rater agreement between the caregiver and the child's rating of the child's oral health-related quality of life (OHRQoL).

Methods

Sample

The sample for this study was a nonrandom, convenience sample of, male and female children, who were recruited at University of Medicine and Dentistry New Jersey (UMDNJ), McGill University Dental School and Montreal Children's Hospital McGill University in Montreal Canada, and New York University (NYU) Dental School and NYU Medical Center in New York City. The sample consisted of children seeking pediatric dental or orthodontic care at UMDNJ, McGill University or NYU and children seeking craniofacial care at McGill University or NYU. Trained interviewers approached individuals scheduled for appointments who were identified, whenever possible, in advance as being in the correct age range for the study. Children between the ages of 8 and 15 years whose caregiver spoke either English or Spanish or French were recruited. Children were excluded if they were enrolled in a special class for mental disabilities or had a mental disorder. Presence of mental disorder was determined by the clinic staff at sites in NJ and NY and information was also solicited from caregivers. After describing the purpose of the study, participating children assented and caregivers consented to a protocol approved by the Institutional Review Boards at the respective locations. Participants independently completed the COHIP and received token payment for their time and effort.

Measures

Demographic Data

Each child participant was asked to report his or her age in years, race/ethnicity, grade in school, and gender. For ethnicity the eight choices presented were Asian, East Indian, Latino, Black, White, Native American, Other and none.

COHIP

The Child Oral Health Impact Profile (COHIP) consists of 34 items forming five conceptually distinct subscales: *Oral Health, Functional Well-being, Social-Emotional Well-being, School-environment and Self-image*. The questions are formatted so as to elicit self-reports from the child and proxy reports from the caregivers (16). Scores can range from 0–136 for the overall scale. Subscales scores are calculated by summing the responses of the items specific to the subscale. The overall OHRQoL score are computed by summing the subscales scores.

Statistical analysis

Spearman correlation

Spearman correlation is a nonparametric measure of correlation computed without making any assumptions about the frequency distribution of the variables. It is the correlation between the data that have been ranked and is appropriate when data may be skewed and when there may be outliers. The Spearman correlations were considered in terms of conventional levels of significance ($P < .05$).

Intra-class correlation

The intra-class correlation is a measure of the proportion of overall variability accounted for by the variability among individuals. High ICC suggests little variability of measurement by either caregiver or child (e.g. the ratings may be interchangeable). Intra-class correlation coefficients are considered excellent if greater than 0.74, good if ranging from 0.60 to 0.74, and fair if ranging from 0.40 to 0.59 (17).

Discrepancy scores

Discrepancy scores for each subscale and the overall COHIP score were computed for each dyad. The child score was subtracted from the caregiver score yielding a difference score that could either be negative (higher child report of QoL) or positive (higher caregiver report of QoL). The differences for each subscale and the overall score were then categorized separately using the standard deviation of the dyad's (child and caregiver) difference scores for the specific subscale or overall score (11). Dyads with raw discrepancy scores ranging from one-half standard deviation above to one-half standard deviation below zero were classified as having similar reports (Similar). Those with raw difference scores greater than one-half standard deviation below zero were classified as children reporting greater QoL than the caregiver did (child more positive). Those with raw difference scores greater than one-half a standard deviation above zero were classified as caregivers reporting greater QoL than the child (child more negative).

Kruskal–Wallis testing

The Kruskal–Wallis test (or H test) is a nonparametric test of the null hypothesis that the distribution of the response variable is the same in three or more groups. The response variable, here categories of agreement between child and caregiver, must

Table 1. Percent of dyads in each clinical group in each concordance category

COHIP score relative to caregiver	Pediatric (%)	Orthodontic (%)	Craniofacial (%)
Oral health			
Child higher	23.9	30.8	44.9
Similar	30.0	25.6	25.5
Child lower	46.2	43.6	29.6
Functional Well-being			
Child higher	35.0	35.7	49.5
Similar	30.7	34.1	18.6
Child lower	34.3	30.2	32.0
Social-emotional well-being			
Child higher	24.5	36.2	43.9
Similar	36.0	33.9	30.6
Child lower	39.6	29.9	25.1
School environment			
Child higher	15.8	19.2	33.0
Similar	54.7	58.5	40.2
Child lower	29.5	22.3	26.8
Self-image			
Child higher	24.8	33.9	34.7
Similar	24.8	28.4	17.9
Child lower	50.4	37.8	47.4
Overall QoL			
Child higher	23.3	33.0	47.7
Similar	32.5	27.7	17.1
Child lower	44.2	39.3	35.2

Categories created using standard deviation of the dyad's (child and caregiver) difference scores for the specific subscale or overall score. Dyads with raw discrepancy scores ranging from one-half standard deviation above to one-half standard deviation below zero were classified as having similar reports (Similar). Those with raw difference scores greater than one-half standard deviation below zero were classified as children reporting greater QoL than the caregiver did (child more positive). Those with raw difference scores greater than one-half a standard deviation above zero were classified as caregivers reporting greater QoL than the child (child more negative).

be ordinally scaled (Table 1) (18). With a large enough sample, as in the present study, the statistic is distributed as chi-square with $s-1$ degrees of freedom (s being the number of groups). We report the Kruskal–Wallis statistic for row mean scores which indicates whether there is a difference in the variable if interest between the groups under consideration (e.g. pediatric, orthodontic, craniofacial) (19). Kruskal–Wallis testing was done comparing the concordance categories and clinical groups (3 × 3) for the overall COHIP score as well as the individual subscales. Kruskal–Wallis testing was also completed to compare concordance categories and child gender (3 × 2), and child age (3 × 2, less than 12 years and 12 years and older, median split), child ethnicity (3 × 4, Black, Latino, White, Other) and location (3 × 3) for the overall COHIP score as well as the individual subscales.

When the obtained value of the H-statistic is statistically significant, it indicates that at least one of the groups is different from the others. It does not indicate, however, which groups are different or where the differences lie. The next step is to construct pair-wise multiple comparisons to locate the source of significance. The level of significance set for each comparison must be adjusted appropriately for the number of differences being tested (20). Thus, if the overall H-test is significant the percentage of dyads in the groups of interest are compared using the z-ratio to test for the significance of the difference between two proportions with the P-value adjusted for the number of comparisons. The percentage of 'child higher' to the percentage 'similar' and the percentage 'child lower' to the percentage 'similar' within each group of interest are tested. P-values are adjusted for the number of pair-wise comparisons.

Response rates

The response rate for subject accrual is unknown because information on nonparticipants was not collected.

Results

Sample

437 caregivers and 419 children completed the COHIP questionnaire; complete information was available for 376 dyads. The dyads were distributed amongst the clinical groups as follows: 141 pediatric, 135 orthodontic, and 100 craniofacial. The dyads were distributed at recruitment locations as follows: Montreal (50 pediatric, 13 orthodontic, 15 craniofacial), UMDNJ (45 pediatric, 15 orthodontic, 0 craniofacial), and NYU (46 pediatric, 107 orthodontic, 85 craniofacial).

The ethnicity of the children was 25.1% Latino, 21.4% Black, 43.9% White, and 9.6% other. 50.5% of the children were female. The children's age ranged from 8 to 15 (M = 11.8 years, SD = 1.5) and were attending grades 1–10. The ethnicity of the caregivers was 24.8% Latino, 21.3% Black, 45.1% White, and 8.8% other. 76% of the caregivers were female. The caregiver's ranged in age from 24 to 68 (M = 41.2 years, SD = 8.0; Median = 41.0). The reported educational level of the caregivers averaged 13.3 years (SD = 3.7; Median = 12). Reported caregiver education ranged from virtually no formal education to post-graduate training. Table 2 presents the demographic characteristics of the dyads retained in the sample as well as those dropped. The children who were dropped from the sample were slightly older (M = 12.6, SD = 1.5, $t = -3.3$, $P < .0009$) and thus not surprisingly had more education (M = 5.8, SD = 1.5, $t = -2.4$, $P < .01$). Caregivers from Montreal were more likely to be dropped from the sample ($z = -1.9$, $P < .03$). In all cases, the child or caregiver who participated was dropped because the complementary questionnaire was not completed.

Correlations between caregiver and child reports

Spearman correlations and Intra-class correlations (ICC) between the caregivers' and children's responses to the individual COHIP subscales and the overall COHIP were examined (Table 3). Correlations are presented for the three groups and for the total sample. Using the Spearman correlations for the sample as a whole, there was fair to moderate agreement between the child and the caregiver's responses on the overall COHIP as well as on each COHIP scales. Spearman correlations were lowest in the craniofacial group for overall COHIP as well as on three of the five scales (*Functional, School and Self*). The magnitude of the ICCs was generally consistent with the Spearman correlations, where they differ the ICC was usually lower. The ICCs suggest low to fair agreement between child and caregiver reports.

Rates of agreement and direction of discordance by clinical group

Table 1 presents the percent of dyads in each clinical group where the child rated OHRQoL

Table 2. Demographic characteristics for the dyads retained and for children and caregivers dropped from the sample

	Child retained (n = 376)	Child dropped (n = 43)	Caregiver retained (n = 376)	Caregiver dropped (n = 21)
Gender *n* (%)				
Male	186 (49.5)	22 (51.2)	90 (24.0)	6 (28.6)
Female	190 (50.5)	21 (48.8)	285 (76.0)	15 (71.4)
Race/ethnicity *n* (%)				
Latino	94 (25.1)	14 (32.6)	93 (24.8)	6 (28.6)
Black	80 (21.4)	10 (23.3)	80 (21.3)	1 (4.76)
White	164 (43.9)	16 (37.2)	169 (45.1)	13 (61.9)
Other	36 (9.6)	3 (7.0)	33 (8.8)	1 (4.76)
Education (grade) M (SD)	5.21 (1.6)	5.8 (1.5)	13.3 (3.7)	12.8 (4.2)
Age (years) M (SD)	11.8 (1.5)	12.6 (1.51)	41.2 (8.0)	39.2 (7.0)
Clinical Group *n* (%)				
Pediatric	141 (37.5)	17 (39.5)	141 (37.7)	11 (52.4)
Orthodontic	135 (35.9)	16 (37.2)	135 (35.9)	8 (38.1)
Craniofacial	100 (26.6)	10 (23.3)	100 (26.6)	2 (9.5)
Data collection location *n* (%)				
Montreal	78 (20.7)	7 (16.3)	78 (20.7)	8 (38.1)
New Jersey	60 (16.0)	4 (9.3)	60 (16.0)	3 (14.3)
New York	238 (63.3)	32 (74.4)	238 (63.3)	10 (47.6)

Table 3. Spearman and intraclass correlations between child and caregiver subscales and overall COHIP scores

	Pediatric		Orthodontic		Craniofacial		Total	
	r	ICC	*r*	ICC	*r*	ICC	*r*	ICC
Oral health	.42 d	.35	.24 b	.25	.41 d	.37	.33 d	.32
Functional well-being	.27 b	.25	.38 d	.41	.20 a	.19	.41 d	.41
Social-emotional well-being	.37 d	.30	.43 d	.45	.40 d	.43	.41 d	.41
School environment	.27 b	.21	.32 c	.20	.22	.28	.31 d	.26
Self-image	.28 b	.27	.29 b	.31	.20	.29	.27 d	.29
overall QoL	.48 d	.37	.40 b	.41	.34 b	.34	.43 d	.41

a = $P < .05$; b = $P < .01$; c = $P < .001$; d = $P < .0001$. ICC, intraclass correlation.

higher, who agreed (responses were similar), and the percent where the child rated the OHRQoL lower than the caregiver. Two issues/questions can be addressed by examining these data: (1) Are rates of agreement similar across clinical groups?; and (2) Are parents of children in any of the clinical groups more likely to rate the child's QoL more positively or more negatively than parents in the other clinical groups?. For example, on the *Oral Health* domain in the craniofacial group, 25.5% of the caregiver/child dyads were in the similar category, while for 44.9% of the dyads the child reported more positive OHRQoL than the caregiver. In 29.6% of the dyads, the child reported more negative OHRQoL than the caregiver.

Overall COHIP

The Kruskal–Wallis test indicated that the proportion of discordance amongst the clinical groups was significantly different for the overall COHIP score H (2) = 14.1, $P = .0009$ indicating that there are differences between the three groups for this score. The comparisons after the *H*-test for the overall COHIP (and subsequent subscales) are *z*-tests testing the difference between the percentage of 'child higher' as compared to the percentage 'similar' and between the percentage 'child lower' to the percentage 'similar' within each clinical group. *P*-values are adjusted for the six pair-wise comparisons (0.05/6 = 0.008). Within the craniofacial group the child was more likely to rate QoL higher (47.7%) than the dyad was to give similar scores (17.1%) ($z = 4.35$, $P < .0001$). Likewise within the craniofacial group the child was more likely to rate QoL lower (35.2%) than the dyad was to give similar scores (17.1%) ($z = 2.75$, $P < .003$). *Z*-testing on the same comparisons within the pediatric and orthodontic groups was not significant.

Oral health

The proportion of discordance amongst the clinical groups was significantly different for the *oral health* subscale H (2) = 11.2, $P = .004$ indicating that there are differences between the three groups for this score. Within the craniofacial group the child was more likely to rate *oral health* higher (44.9%) than the dyad was to give similar scores (25.5%) (*z* = 2.84, *P* < .002). Within the pediatric group the child was more likely to rate *oral health* lower (46.2%) than the dyad was to give similar scores (30.0%) (*z* = 2.68, *P* < .004). Within the orthodontic group the child was more likely to rate *oral health* lower (43.6%) than the dyad was to give similar scores (25.6%) (*z* = 3.09, *P* < .001).

Functional well-being

The proportion of discordance amongst the clinical groups was significantly different for the *functional well-being* subscale H (2) = 6.15, *P* = .05 indicating that there are differences between the three groups for this score. Within the craniofacial group the child was more likely to rate QoL higher (49.5%) than the dyad was to give similar scores (18.6%) (*z* = 4.55, *P* < .0001). The child was also more likely to rate QoL lower (32.0%) than the dyad was to give similar scores (18.6%) (*z* = 2.15, *P* = .02) but this difference failed to reach the *P*-value corrected for multiple comparisons. *Z*-testing on the same comparisons within the pediatric and orthodontic groups was not significant.

Social-emotional well-being

The proportion of discordance amongst the clinical groups was significantly different for the *social emotional well-being* subscale H (2) = 10.8, *P* = .005 indicating that there are differences between the three groups for this score. The percentage of 'child higher' to the percentage 'similar' and the percentage 'child lower' to the percentage 'similar' within each clinical group were tested. Although none reached the adjusted *P*-value for significance, within the craniofacial group the child showed a tendency to rate *social emotional well-being* higher (43.9%) as compared to the dyad giving similar scores (30.6%) (*z* = 1.92, *P* < .03). Also not reaching the adjusted level of significance was the comparison within the pediatric group that suggested that the dyad was more likely to report similar scores (36.0%) as compared to the child reporting higher QoL (24.5%) (*z* = −2.09, *P* < .02).

School environment

The proportion of discordance amongst the clinical groups was significantly different for the *school environment* subscale H (2) = 9.4, *P* = .009 indicating that there are differences between the three groups for this score. Within the pediatric group the dyad was more likely to report similar scores (54.7%) as compared to the child reporting higher QoL (15.8%) (*z* = −6.8, *P* < .0001). Also within the pediatric group the dyad was more likely to report similar scores (54.7%) as compared to the child reporting lower QoL (29.5%) (*z* = −4.25, *P* < .0001). Similarly, within the orthodontic group the dyad was more likely to report similar scores (58.5%) as compared to the child reporting higher (19.2%) (*z* = −6.5, *P* < .0001) or lower QoL (22.3%) (*z* = −5.91, *P* < .0001). *Z*-testing on the same comparisons within the craniofacial group did not reach the adjusted level of significance.

Self-image

The proportion of discordance amongst the clinical groups was not significantly different for the *self-image* subscale H (2) = 3.1, *P* = 0.22. No further testing was performed.

Direction of discordance by child gender, age, race/ethnicity and location

Kruskal–Wallis testing was also completed to compare concordance categories and child gender (3 × 2), and child age (3 × 2, less than 12 years and 12 years and older, median split), child ethnicity (3 × 4, Black, Latino, White, Other) and location (3 × 3, Montreal, NJ, NY) for the overall COHIP score as well as the individual subscales. The proportion of discordance amongst the locations was significantly different for the *school environment* subscale and for the *self-image* subscale. All remaining tests on gender, age, race/ethnicity and location were associated with *P*-values > 0.05 (data not shown).

Kruskal–Wallis testing on the *school environment* subscale produced an H (2) = 6.15, *P* = .046 indicating that there were differences between the three locations for this subscale score (Table 4). Within the Montreal location the dyad was more likely to report similar scores (46.7%) as compared to the child reporting higher QoL (16.0%) (*z* = −4.05, *P* < .0001). Within the NJ location the dyad was more likely to report similar scores (54.2%) as compared to the child reporting higher (20.4%) (*z* = 3.81, *P* < .0001) or lower (25.4%) (*z* = 3.2, *P* = .0007) QoL. Similarly, within the NY location the dyad was more likely to report similar scores (53.5%) as compared to the child reporting higher (23.7%) (*z* = 6.58, *P* < .0001) or lower (22.8%) (*z* = 6.8, *P* < .0001) QoL.

Kruskal–Wallis testing on the *self-image* subscale H (2) = 8.4, *P* = .015 indicating that there were

Table 4. Percent of Dyads from each Location in each Concordance Category

COHIP Score relative to Caregiver	Montreal McGill	NJ UMDNJ	NY NYU
School Environment			
Child higher	16.0	20.4	23.7
Similar	46.7	54.2	53.5
Child lower	37.3	25.4	22.8
Self-image			
Child higher	19.8	24.6	35.7
Similar	23.9	26.3	23.8
Child lower	56.3	49.1	40.5

Categories created using standard deviation of the dyad's (child and caregiver) difference scores for the specific subscale or overall score. Dyads with raw discrepancy scores ranging from one-half standard deviation above to one-half standard deviation below zero were classified as having similar reports (Similar). Those with raw difference scores greater than one-half standard deviation below zero were classified as children reporting greater QoL than the caregiver did (child more positive). Those with raw difference scores greater than one-half a standard deviation above zero were classified as caregivers reporting greater QoL than the child (child more negative).

differences between the three locations for this subscale score (Table 4.). Within the Montreal location, the child was more likely to rate QoL lower (56.3%) than the dyad was to give similar scores (23.9%) ($z = 3.94$, $P < .0001$). Likewise within the NJ location the child was more likely to rate QoL lower (49.1%) than the dyad was to give similar scores (26.3%) ($z = 2.51$, $P = .006$).

Within the NY location the child was more likely to rate QoL higher (35.7%) ($z = 2.77$, $P = .003$) or lower QoL lower (40.5%) ($z = 6.8$, $P < .0001$) than the dyad was to give similar scores (23.8%).

Given the location findings, the Kruskal–Wallis comparisons of the clinical groups were repeated within location for the *school environment* and *self-image* subscales. Because cell sizes were smaller within location with a number having counts less than 5, Fisher's exact test was also computed. Comparison of the *school environment* subscale scores between clinical groups was significant for the NY location H (2) = 10.9, $P = .004$ only. Comparison of the *self-image* subscale scores between clinical groups within location yielded no significant results.

Discussion

Low to modest rates of agreement between child and caregiver were found for the sample overall. Rates of agreement did vary between clinical groups. The pediatric and orthodontic dyads were more likely to agree and disagree in a similar manner to each other while the pattern in the craniofacial dyads was different. On the *oral health* subscale children in both the pediatric and orthodontic groups were more likely to report lower QoL than they were to agree with the caregiver. In contrast, for not only the *oral health* subscale but on the *functional well-being* and Overall COHIP score children in craniofacial group were more likely to report higher QoL than they were to agree with the caregiver. This tendency was present for the craniofacial children on the *social emotional well-being* subscale but it failed to reach the significance level which had been adjusted for multiple comparisons.

Rates of agreement also varied between clinical groups on the *school environment* subscale with pediatric and orthodontic dyads exhibiting a similar pattern. For both groups the dyads were more likely to agree with one another than they were to disagree in either direction (child higher or child lower). The craniofacial group did not evidence a similar pattern with no significant differences between categories of concordance detected. These results should be viewed cautiously because location differences were found for the *school environment* subscale. While no differences between concordance categories were found for clinical groups in Montreal or New Jersey, there were significant differences for the dyads from NY. The pattern of the differences in NY was the same as the finding in the overall sample (dyads more likely to agree). It is possible that the samples from Montreal and NJ were too small to detect significant differences between groups when school environment was examined within location. It is also important to note that the majority of craniofacial dyads came from NY and they did not follow the same pattern (dyads more likely to agree). Furthermore with the large number of comparisons made (age, gender, ethnicity, location) these findings may be spurious. It will be essential in continuing this research that adequate samples be captured from data collection locations to explore this finding further.

We were unable to detect differences in concordance between caregiver and child on ratings of the child's OHRQoL that were associated with the child's age, gender or ethnicity. This finding is consistent with other reports (21, 22). Furthermore, when associations are found the results are not generally straight forward or easy to interpret (4). It is important to continue to consider the role such demographic characteristics might play in assessing concordance both at an aggregate and individual level so as not to miss the role they might play because it is not easily determined.

The direction of the discrepancies for the craniofacial dyads was consistent with children reporting higher OHRQoL than the caregiver in each instance where they disagreed, including the overall OHRQoL rating. Because many craniofacial anomalies are present from birth, caregivers of these children have been living with and taking care of a child who has always had health concerns. Thus the chronic nature of the condition has been found to manifest in parental over-protectiveness of the affected child and could impact their perceptions of the child's quality of life (9, 23). Furthermore, caregivers of craniofacial patients reportedly have higher parental stress than parents having children with no congenital malformations (24, 25). Parental mental health has been linked to altered ratings of the child's QoL, specifically with mothers who had higher distress levels rated their child's HRQoL lower than the child rated it (11).

It is notable that we were unable to detect significant differences among clinical groups for rates of concordance for the self-image subscale. Given the fact that significant differences were detected on all other subscales and on the Overall COHIP, this null finding compels us to consider it closely. Inspection of the data indicates that the self-image subscale had the lowest Spearman's correlation coefficient for the total sample (Table 3). It is possible that the apparent the lack of difference among clinical groups is a result of greater disagreement, overall, between children and parents with regard to these questions on self-image.

The results, although more robust for the craniofacial group than the pediatric or orthodontic groups, suggest that when it is at all possible, children should be asked to provide self-ratings of OHRQoL in addition to proxy ratings. This approach is consonant with the idea that information from multiple informants can be used to provide a well-rounded depiction of the child and his or her health care needs and personal quality of life issues and concerns. For example, the UK National Service Framework for Children now requires services to proactively elicit and act upon the views of children and their care providers, in order to involve children in their healthcare (2, 26) and thereby facilitate child-centered health care.

Child-centered care is particularly relevant for youth with craniofacial anomalies where treatment (e.g. lip/nose revisions, pharyngeal flaps, bone grafts) is often painful, time-consuming and expensive events requiring hospitalization and possible subsequent therapy. A 1997 (27) study examined the psychological outcomes of cleft treatment in not only the late adolescent patients (aged 15–20) but also the parents' assessment of the treatment. A small but significant percentage of the youngest participants felt excluded from treatment decisions. No agreement between caregivers and patients regarding their satisfaction with clinical outcome was found – leading the authors to conclude that young patients should be queried independently on planned treatment for such conditions. If for example, the parental opinion underestimates the child's self-reported OHRQoL and the underestimate is guiding treatment decisions, the child's best interests may not be best served. Further in the case of treatment requiring subsequent therapy by multiple specialists (e.g. speech), adherence with regimens could be jeopardized without a commitment from both caregivers and the children.

It is a limitation of this study that the sample was not balanced in terms of clinical groups being drawn equally from the different data collection sites. The majority of craniofacial dyads came from NY and far more orthodontic patients were from the NY location than from Montreal or NJ locations as well. When site differences were tested for the five subscales and the Overall COHIP score, only the school subscale evidenced differences by site. Further exploration of this finding revealed that no differences were found for clinical groups in Montreal or New Jersey but significant differences for the dyads from NY detected. The possible implications of this result have been discussed previously; it is essential that future work capture adequate samples from varoius clinical conditions as well as balanced representation from data collection sites.

In summary, the present study found that rates of concordance between child and caregiver when rating the child's OHRQoL varied between clinical groups. Patients with craniofacial anomalies were less likely to agree with their caregivers' and were more likely to rate their OHRQoL more positively. Although the finding was more limited, pediatric and orthodontic patients were more likely to agree with their caregivers' or to rate their OHRQoL more negatively. These findings on concordance using the COHIP provide further evidence of the usefulness of both child and caregiver reports.

Acknowledgement

This project was supported by the NIDCR: R01-DE13722.

References

1. Wallander JL, Schmitt M, Koot HM. Quality of life measurement in children and adolescents: issues, instruments, and applications. J Clin Psychol 2001;57:571–85.
2. Vogels T, Verrips GHW, Verloove-Vanhorick SP, Fekkes M, Kamphuis RP, Koopman HM et al. Measuring health-related quality of life in children: the development of the TACQOL Parent Form. Qual Life Res 1998;7:457–65.
3. Eiser C, Morse R. Review of measures of quality of life for children with chronic illness. Arch Dis Child. 2001;84:205–11.
4. Eiser C, Morse R. Quality-of-life measures in chronic diseases of childhood. UK: In Health Technology Assessment; 2001. p. 1–156.
5. Achenbach TM, McConanghy SH, Howell CT. Child-adolescent behavioral and emotional problems: implications ok cross-informant correlations for situational specificity. Psychol Bull 1987;101:213–32.
6. Bates JE, Pettit GS, Dodge KA, Ridge B. Interaction of temperamental resistance to control and restrictive parenting in the development of externalizing behavior. Dev Psychol 1998;34:982–95.
7. Theunissen NCM, Vogels TGC, Koopman HM, Verrips GHW, Zwinderman KAH, Verloove-Vanhorick SP et al. The proxy problem: child report versus parent report in health-related quality of life research. Qual Life Res 1998;7:387–97.
8. Ennett ST, DeVellis BM, Earp JA, Kredich D, Warren RW, Wilhelm CL. Disease experience and psychosocial adjustment in children with juvenile rheumatoid arthritis: children's versus mothers' reports. J Pediatr Psych 1991;16:557–68.
9. Broder HL, Smith FB, Strauss RP. Developing a behavioral rating scale for comparing teachers' rating of children with and without craniofacial anomalies. Cleft Palate/Cranio J. 2001;38:560–6.
10. Levi R, Drotar D. Critical issues and needs in health-related quality of life assessment of children and adolescents with chronic health conditions. In: Drotar D editor. Measuring health-related quality of life in children and adolescents: Implications for research and practice. Mahwah, N.J.: Lawrence Erlbaum Associates; 1998. p. 3–25.
11. Garber J, Van Slyke DA, Walker LS. Concordance between mothers' and children's reports of somatic and emotional symptoms in patients with recurrent abdominal pain or emotional disorders. J Abnorm Child Psychol 1998;26:381–91.
12. Slade GD. The Oral Health Impact Profile. In: Slade GD editor. Measuring Oral Health and Quality of Life. North Carolina: Univ of North Carolina Press; 1997. p. 93–104.
13. Inglehart M, Bagramian RA. Oral Health-related Quality of Life. Chicago: Quintessence Publishing; 2002.
14. Jokovic A, Locker D, Stephens M, Kenny D, Tompson B, Guyatt G. Validity and reliability of a questionnaire for measuring child oral-health-related quality of life. J Dent Res 2002;81:459–63.
15. Broder HL, Wilson-Genderson M. Reliability and validity of the Child Oral Health Impact Profile (COHIP). Community Dent Oral Epidemiol 2007; 34(Suppl. 1):20–31.
16. Fleiss J. Statistical methods for rates and proportions. 2nd ed. New York, NY: John Wiley & Sons; 1981.
17. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. J Am Stat Assoc 1952;47:583–621.
18. Stokes M, Davis C, Koch GG. Categorical Data Analysis using the SAS System. Cary NC: SAS Institute; 1995.
19. Hettmansperger TP. Statistical Inference Based on Ranks. New York, NY: John Wiley & Sons; 1984.
20. Barrett ML, Berney TP, Bhate S, Famuyiwa OO, Fundudis T, Kolvin I et al. Diagnosing childhood depression who should be interviewed–parent or child? The Newcastle Child Depression Project Br J Psychiatry Suppl 1991;11:22–7.
21. McConaughy SH, Stranger C, Achenbach TM. Three-year course of behavior/emotional problems in a national sample of 4-to 16-year-olds, I: agreement among informants. J Am Acad Child Adolesc Psychiatry 1992;31:932–40.
22. Tobiasen JM, Levy J, Carpenter MA, Hiebert JM. Type of facial cleft, associated congenital malformations, and parents' ratings of school and conduct problems. Cleft Palate J 1987;24:209–15.
23. Abidin RR, Wilfong E. Parenting stress and its relationship to child health care. Child Health Care 1989;18:114–6.
24. Pope AW, Speltz ML. Research of psychosocial issues of children with craniofacial anomalies: progress and challenges. Cleft Palate Craniofac J 1997;34:371–3.
25. Kuther TL. Medical decision-making and minors: issues of consent and assent. Adolescence 2003;38:343–58.
26. Turner SR, Thomas PW, Dowell T, Rumsey N, Sandy JR. Psychological outcomes amongst cleft patients and their families. Br J Plast Surg 1997;50:1–9.