


# Who Should Fill Out a Pediatric PROM? Psychometric Assessment From a Clinical Perspective in 567 Children With a Cleft

FACE  
2024, Vol. 5(1) 50–58  
© The Author(s) 2023  
  
DOI: 10.1177/27325016231209051  
journals.sagepub.com/home/fac



Saranda Ombashi<sup>1,2</sup> , Victor L. van Roey<sup>1,2</sup>,  
Jolanda M. E. Okkerse<sup>1</sup>, Mariska van Veen-van der Hoek<sup>1,3</sup>,  
Esther E. B. van Oers-Hazelzet<sup>4</sup> ,  
Aebele B. Mink van der Molen<sup>2,4</sup>, and Sarah L. Versnel<sup>1,2,3</sup>

## Abstract

**Background:** The CLEFT-Q is a questionnaire developed for patients with a cleft lip and/or palate (CL/P). Numerous scales have been implemented as part of the ICHOM Standard Set for CL/P. Although validated for completion by patients only, clinicians noted that caregivers are often involved in completion of the scales. Aim of the study was to promote further standardization of Patient Reported Outcome Measures (PROMs) in pediatric patients by examining the preferences of patients and parents concerning the reporter type. Moreover, possible discrepancies in outcomes between reporter types were explored. **Methods:** Data from 567 patients with CL/P and their caregivers that completed scales of the CLEFT-Q questionnaire were collected. Reporter group sizes and proportions were examined at the ages of 8, 12, and 15 years to determine the preferred manner of completion. Mean outcomes were analyzed per scale at the 3 ages, and compared between the 3 reporter groups: “patient,” “caregiver,” and “together.” **Results:** In all age-groups, the majority completed the PROMs together. Concerning the reporter types per age-group, an upward trend was seen in the proportion of patients that completed the scales alone. In the caregiver group, a downward trend was observed, and the highest proportion of parents that completed the scales was found at age 8. No significant differences were found between the reporter types in any of the scales. **Conclusion:** Even if a PROM questionnaire is validated for patient report only, it is recommended to record the reporter type when a pediatric PROM is completed. In order to capture outcomes that represent the patient’s voice validly and reliably, though with support of the caregiver, a pediatric PROM should be filled out by the patient alone and thereafter evaluated with the caregiver(s). Concerning the CLEFT-Q, there seems to be demand for a validated parent-version of the scales.

## Keywords

cleft lip and palate, patient reported outcome measures, standardized outcome measures, pediatric quality of life, facial appearance

## Introduction

In the last decades, systematic outcome measurement has become a fundamental condition for value-based healthcare, and quality improvement in pediatric healthcare seems to increasingly depend on the standardization of these outcome measurements.<sup>1,2</sup> Patient-Reported outcome measures (PROMs) play a key role in systematic outcome measurement, as they provide unique information of that what matters most to patients.<sup>3,4</sup>

Concerning cleft care, many calls and large initiatives in order to find or develop standardized outcome measures were set up over the past decennia, such as Eurocleft, Scandcleft, Americleft and Eurocat.<sup>5-9</sup> Unfortunately, no validated

PROMs for children with a cleft lip and/or cleft palate (CL/P) were used in the initiatives. However, in 2017, the International

<sup>1</sup>Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

<sup>2</sup>European Reference Network for Rare and/or Complex Craniofacial Anomalies and Ear, Nose, and Throat Disorders

<sup>3</sup>Erasmus University Medical Center, Rotterdam, The Netherlands

<sup>4</sup>UMC Utrecht, University Medical Utrecht, Utrecht, The Netherlands

### Corresponding Author:

Saranda Ombashi, Department of Plastic and Reconstructive Surgery, Erasmus MC, University Medical Center Rotterdam, Erasmus University Medical Center, Dr. Molewaterplein 40, Rotterdam 3015 GD, The Netherlands.

Email: s.ombashi@erasmusmc.nl

Consortium for Health Outcomes Measurement (ICHOM) developed the ICHOM Standard Set for Cleft Lip and Palate (ICHOM Standard Set).<sup>10</sup> An effort was taken to find a multi-disciplinary, international consensus for a standard set of outcomes in cleft care.<sup>10</sup> Concerning PROMs, the workgroup found consensus in implementing several scales of the CLEFT-Q questionnaire into the ICHOM Standard set.<sup>10</sup>

The CLEFT-Q questionnaire is a cross-cultural PROM, made recently, that is specifically developed and validated for patients with CL/P.<sup>11-13</sup> Erasmus Medical Center (EMC, Rotterdam, The Netherlands), and University Medical Center Utrecht (UMCU, Utrecht, The Netherlands) have both implemented the ICHOM Standard Set. It was noticed that in clinical practice, caregivers were often involved in completing the CLEFT-Q. However, the questionnaire is only validated for completion by patients. And although psychometric properties of the CLEFT-Q have been examined thoroughly in previous studies, no studies regarding the effect of reporter have been done so far.<sup>13,14</sup>

Other pediatric PROM instruments, such as the PedsQL, have reported discrepancies between patient- and parent-reported outcomes in their studies.<sup>15-17</sup> Therefore, many validated generic pediatric PROM instruments are available with scales specifically for patients or parents, or 2 completely separate versions of a questionnaire are used.<sup>18-25</sup> More child-parent agreement in physical domains than in social and emotional domains is reported, although other studies state that agreement depends on the clinical relevance to a given disease group.<sup>15,16</sup>

In order to encourage further standardization of outcome measures and to enable future research and benchmarking, it is essential to examine to what extent reporter type affects CLEFT-Q outcomes.

The objective of this multicenter study is to promote further standardization of pediatric PROMs by exploring possible discrepancies between reporter types in children with a cleft, using all CLEFT-Q scales included in the current ICHOM Standard set. This was done by comparing outcome scores between reporter types. Since no randomization methods were applied, the reporter group sizes provide additional insights into the preferred manner of filling out the PROMs for pediatric patients and their caregivers.

## Methods

A retrospective study was conducted on patients with CL/P at the EMC and UMCU using the ICHOM Standard set. The set was implemented in 2015 and 2021 at EMC and UMCU respectively, as part of the regular care protocol for patients with CL/P (ethical board approval nr. METC395478). According to the set, patients are evaluated from their initial visit to the cleft team until they are 22 years old, with the first visit usually occurring during the neonatal period. However, patients who are adopted or patients that have switched cleft

teams are included too, and might be older at the start of treatment in the cleft centers. The ICHOM Standard set comprises various scales of the CLEFT-Q questionnaire, and the assessment of these scales depends on the patient's cleft type and is conducted at the ages of eight, 12, and 22 years.<sup>10</sup> In the UMCU, an additional assessment at age 15 is conducted as well.

## Cleft-Q Questionnaire

The CLEFT-Q questionnaire is a PROM designed specifically for individuals with CL/P. It is internationally applicable as it is available in various languages and has a cross-cultural character.<sup>14,26</sup> The questionnaire consists of 12 scales and one checklist that cover the domains of facial function, health-related quality of life (HRQoL), and appearance. These scales can be used independently or in combination with other CLEFT-Q scales to gain insights into the Quality of Life (QoL) of individuals with CL/P. The ICHOM Standard set includes 8 CLEFT-Q scales and one eating and drinking checklist, including one functional scale (CLEFT-Q Speech Function), 3 appearance scales (CLEFT-Q Face, Teeth, Jaw), and 4 HRQoL scales that evaluate psychological and social domains (CLEFT-Q Speech distress, School function, Social function, Psychological function). The development and validation process of the CLEFT-Q questionnaire are extensively described elsewhere.<sup>11,12,14,27</sup>

## Patient Population

The study included patients with CL/P and their caregivers who were evaluated according to the ICHOM Standard set between 2015 and 2021, for the EMC and UMCU respectively, until January 2023. Cases with at least one completed CLEFT-Q scale and registration of the reporter type were analyzed. Patients with cognitive impairments that prevented them from understanding the questionnaire items were not assessed with any PROM scales of the ICHOM Standard set. The patient's cognition was estimated by the specialized nurse of the cleft team in consultation with the patient's caregivers. Patients who were not proficient in the Dutch language were not assessed either.

## Data Collection

Data collection was part of the regular care protocol, and patients could complete the CLEFT-Q scales either at home by receiving them via email, or during a regular visit to the cleft team. At the outpatient clinic, patients and their caregivers were asked who responded to the CLEFT-Q items, and the responses were categorized as follows: (1) "Patient" meaning the patient completed all items alone; (2) "Caregiver" referring to the parent/caregiver(s) who responded to all items without the involvement of the patient; (3) "Together" indicating the patient filled out the

**Table 1.** Baseline Characteristics of the Included Patients, Per Age.

	8y (N=268)	12y (N=243)	15y (N=56)	Overall (N=567)
Sex				
Male	161 (60.1%)	153 (63.0%)	33 (58.9%)	347 (61.2%)
Female	107 (39.9%)	90 (37.0%)	23 (41.1%)	220 (38.8%)
Cleft type				
CL	28 (10.4%)	19 (7.8%)	4 (7.1%)	51 (9.0%)
CLA	29 (10.8%)	22 (9.1%)	7 (12.5%)	58 (10.2%)
CLAP	121 (45.1%)	127 (52.3%)	30 (53.6%)	278 (49.0%)
CP	90 (33.6%)	75 (30.9%)	15 (26.8%)	180 (31.7%)
Center				
EMC	178 (66.4%)	176 (72.4%)	0 (0%)	354 (62.4%)
UMCU	90 (33.6%)	67 (27.6%)	56 (100%)	213 (37.6%)

questionnaire with one or 2 caregiver(s). In this study, both (biological) parents and caregivers will be referred to as “caregivers.”

Baseline characteristics of all patients were recorded, including sex and cleft type. Data was collected from the following moments of assessment: Time-point 8 years, 12 years, and 15 years (only UMCU). According to the measurement moments of the ICHOM Standard set, 8-year-olds were evaluated with the CLEFT-Q Face, Teeth, and Social function. In patients who were 12 and 15 years old, assessment took place with the CLEFT-Q Face, Teeth, Jaw, Psychological function, School function, Speech function, and Speech distress scale. Time-point 22 years was not included in the data collection, because the responder type was not recorded in the 22-year-olds. During the data collection period, some scales of the CLEFT-Q were revised: Outcome scores of the older version of the CLEFT-Q Social scale were not used for this study due to the addition of new items.

### Data Analysis

The data was analyzed using R statistical software. The scales were previously validated according to Rasch measurement theory, which showed reliability and validity in all scales except for the Eating and Drinking checklist. Consequently, outcomes of this checklist have to be evaluated per item instead of using the total score per scale, and therefore the checklist was not included in the current study.<sup>13</sup> For the included scales, scores on a 0 to 100 derived from the logit scores were used, with a higher score indicating a better outcome. The normality of distributions of scores was checked first using skewness and kurtosis values. Values greater than +2 or lower than -2 were considered indicative of non-normality.<sup>28</sup> Levene’s tests were done to test the homogeneity of variances. To investigate possible differences between the scores of the 3 reporter groups for all scales separately at each time point, Kruskal-Wallis H tests were subsequently performed. In case of a significant test,

multiple pairwise comparisons were performed using Dunn’s test while applying Bonferroni correction. Moreover, mean outcomes were visualized in bar plots, including the 2.5%-97.5% quantiles to map out response distributions for each scale. Finally, responder distributions were compared between each age to examine possible trends in the responder preference.

## Results

### Baseline Descriptives and General Findings

A total of 567 patients were included in the study, with the majority being male (61.2%, Table 1).

The distribution of cleft types followed the epidemiology of CL/P, with most patients having a CLAP (47.6%), followed by CP (31.7%), CLA (10.2%), and CL (9.0%).<sup>29</sup> Of the included patients, 268 were 8 years old at the time of assessment, 243 were 12 years old, and 56 were 15 years old.

A detailed overview of reporter types, age-groups, and cleft types is provided in Figure 1.

In the majority of the cases, the CLEFT-Q scales were completed by patient and caregiver(s) together (n=440), followed by the caregiver alone (n=81) and the patient alone (n=46).

### Detecting Differences Per Age Group

All distributions of CLEFT-Q scores were considered non-normal based on the skewness and kurtosis values. Hence, the non-parametric Kruskal Wallis H test was used to test for differences on all scales. None of the scales showed significant differences between the reporter types in any of the age-groups (app. 1).

**Scale outcomes at eight years.** Two appearance scales and one psychosocial scale were completed at 8 years (Table 2). The CLEFT-Q Face was completed 268 times, whereas the Teeth and Social function scale were filled out 267 times.

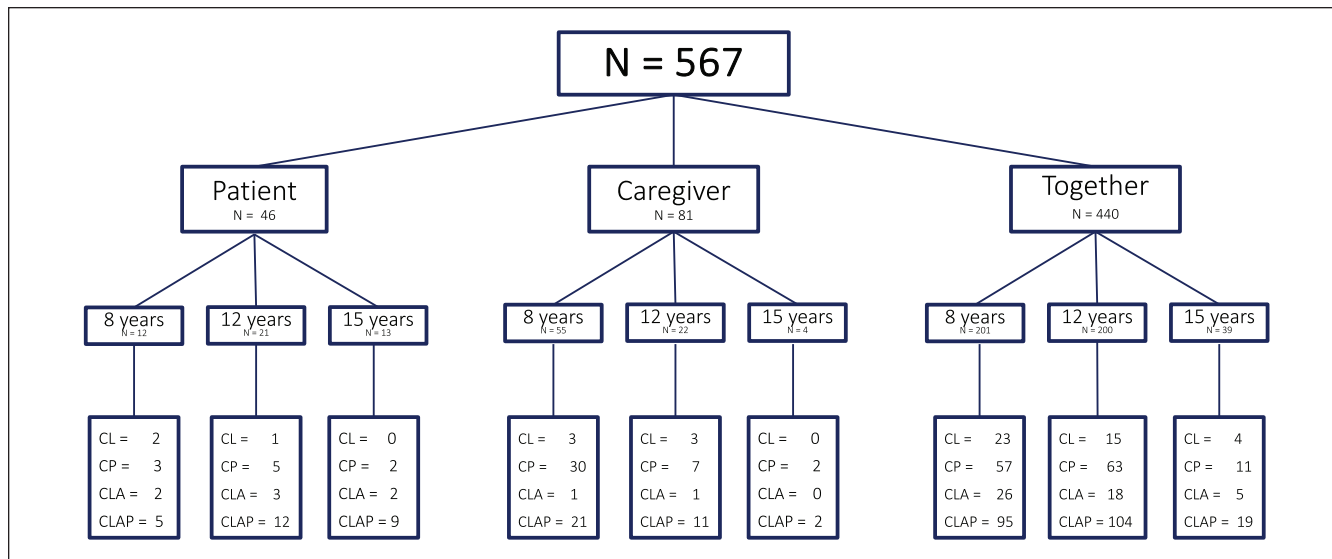


Figure 1. Overview of the participants per responder type.

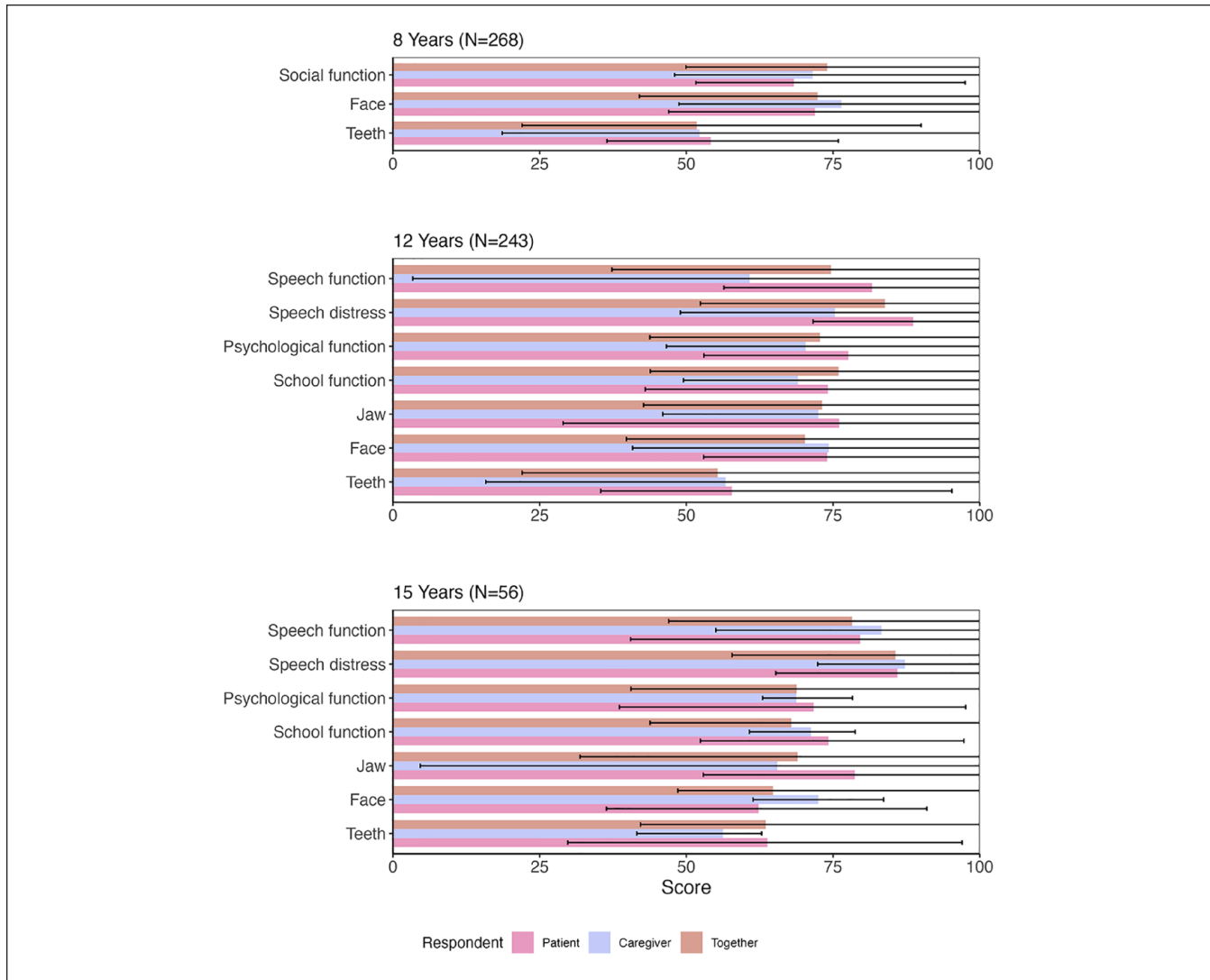
Table 2. Mean Outcomes and 2.5% to 97.5% Quantiles of the CLEFT-Q Scales in Each Age, Categorized by Responder Type.

8y	Patient (N= 12, 4.5%)	Caregiver (N=55, 20.5%)	Together (N=201, 75%)	Overall (N=268)
Face	71.9 (47-100)	76.4 (49-100)	72.4 (42-100)	73.2 (44-100)
Teeth	54.2 (36-76)	52.3 (19-100)	51.8 (22-90)	52.0 (22-100)
Social function	68.3 (52-98)	71.5 (48-100)	74.0 (50-100)	73.2 (48-100)
12y	(N=21, 8.6%)	(N=22, 9.1%)	(N=200, 82.3)	(N=243)
Appearance scales				
Face	74.0 (53-100)	74.3 (41-100)	70.2 (40-100)	70.9 (22-100)
Teeth	57.8 (35-95)	56.7 (16-100)	55.3 (22-100)	55.6 (40-100)
Jaw	76.0 (29-100)	72.5 (46-100)	73.1 (43-100)	73.3 (43-100)
Health-related QoL scales				
School function	74.1 (43-100)	69.0 (50-100)	76.0 (44-100)	75.2 (44-100)
Psychological function	77.6 (53-100)	70.3 (47-100)	72.8 (44-100)	73.0 (44-100)
Speech distress (N=202) <sup>a</sup>	88.6 (72-100)	75.3 (49-100)	83.9 (52-100)	83.5 (52-100)
Functional scale				
Speech function (N=202) <sup>a</sup>	81.6 (56-100)	60.8 (3-100)	74.7 (37-100)	74.0 (34-100)
15y	(N=13, 23.2%)	(N=4, 7.1%)	(N=39, 69.6%)	(N=56)
Appearance scales				
Face	62.3 (36-91)	72.5 (61-84)	64.8 (49-100)	64.8 (43-100)
Teeth	63.8 (30-97)	56.3 (42-63)	63.5 (42-100)	63.1 (32-100)
Jaw	78.7 (53-100)	65.5 (5-100)	68.9 (32-100)	71.0 (19-100)
Health-related QoL scales				
School functioning	74.2 (52-97)	71.3 (61-79)	67.9 (44-100)	69.6 (46-100)
Psychological function	71.7 (39-98)	68.8 (63-78)	68.8 (41-100)	69.5 (35-100)
Speech distress (N=45) <sup>a</sup>	86.0 (65-100)	87.3 (72-100)	85.6 (58-100)	85.9 (60-100)
Functional scale				
Speech function (N=45) <sup>a</sup>	79.6 (41-100)	83.3 (55-100)	78.3 (47-100)	79.0 (47-100)

<sup>a</sup>Only CLAP and CP.

From the reporter types, the “patient alone” formed the smallest group, followed by “parent alone.” The majority completed the scales together (75%, Table 2). In Table 2, an

overview of all mean outcomes and the 2.5% to 97.5% quantiles are provided. No clear trends can be observed: the reporter type with the highest mean score differs in each scale (Figure 2).



**Figure 2.** Bar plots with the outcomes of each scale, categorized per age and responder type. Mean outcomes and 2.5% to 97.5% quantiles are visualized.  
 Note. Appearance scales: Face, Teeth, Jaw. Health-related Quality of life scales: Speech distress, School function, Social function. Functional scale: Speech function.

*Scale outcomes at 12 years.* The Face, Teeth, and Jaw scales were completed 240, 242, and 241 times respectively. The School and Social functioning scales were both filled out 241 times. The CLEFT-Q Speech function and Speech distress were completed in 200 participants. Observing the mean outcome score, the “patients” shows the best scores on most scales. All appearance scales are reported most positively by the “patient” group, although the Face scale shows similar outcomes in “patients” and “caregivers” (Figure 2). In the HRQoL scales, the highest mean score is found in the “patient” group in 2 out of the 3 scales. For the speech function scale, “patients” are the best-scoring group too (Figure 2).

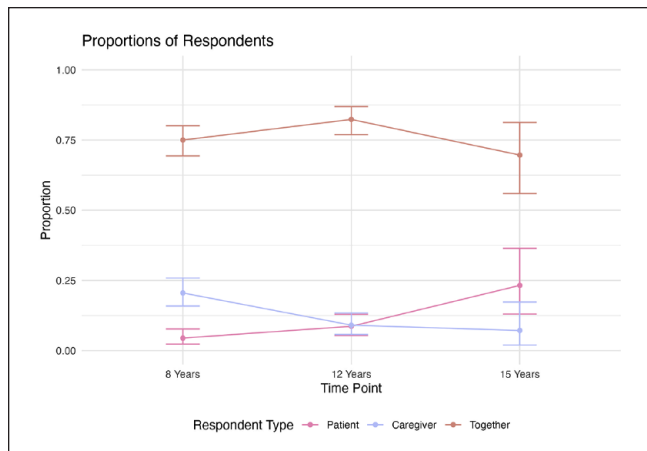
*Scale outcomes at 15 years.* All scales were completed in 56 participants, except for Speech function and Speech distress,

that were completed 45 times. In 2 out of the 3 appearance scales “patients” reported the highest mean scores (Face, Teeth, Figure 2). In the HRQoL scales, the highest mean scores were also found in the “patient” group in 2 out of the 3 scales (Psychological function, Social function, Figure 2). In the Speech function scale, the highest mean score came from the “caregiver” group.

### Detecting Trends in Reporter Types

In Figure 3, trends for the reporter types are given by showing the proportion of participants per responder group at each age.

The confidence intervals per age group and responder type are given in appendix 2.



**Figure 3.** Trend analyses between the age, per responder type. Confidence intervals are visualized as well. Note. X-axis: age of patient at the moment of assessment. Y-axis: proportion of participants per age.

**Table 3.** Two-Sample Test for Equality of Proportions Without Continuity Correction.

	Patient (P)	Caregiver (P)	Together (P)
8vs 12y	.06	<.001	
8vs 15y	<.001	.02	
12vs 15y	<.01	.65	

When focusing on the “patient” group, the largest absolute number is found at age-group 12 years. However, looking at the relative amounts per age group, 15-year olds tend to complete the CLEFT-Q the most often alone (23.2% in 15-year olds, and 8.5% in 12-year olds, Table 2).

In Figure 3, an increasing trend in the “patient” group is visible. The proportion of patients that complete the scales alone increases significantly in 15-year-olds, when compared to 8-year-olds and 12-years-olds (Table 3).

The opposite phenomenon is seen in the “caregiver” group, where a downward trend is observed(Figure 3). The proportion of caregivers that complete the scales is significantly lower in the age-group of 12 years, when compared with the age-group of 8 years.

Overall, the majority of the participants completed the CLEFT-Q together, regardless of the age of the patient. The peak is seen at the 12-year olds (82.3%, Table 2), and declines at age 15, when the “patient” group increases (Figure 3).

**Discussion and Clinical Interpretation**

In this multicenter study of 567 participants, outcomes of all CLEFT-Q scales included in the current ICHOM Standard set were examined to explore if any trends in reporter types are present at the ages of assessment. Possible discrepancies between reporter types were detected at the same ages as

well. This way, preferences of patients and parents were mapped out, and new insights for further improvement of standardization of pediatric PROMs were provided.

**Strengths and Limitations**

This is the first study to explore possible discrepancies in CLEFT-Q scales, based on type of reporter, and therewith provides unique insights in and between these groups. As patients with CL/P and caregivers included in the study individually decided who should complete the questionnaire, both advantages and disadvantages of the study design can be listed. Due to small sample sizes in the “patient” and “caregiver” group, possibilities for statistical analysis are limited in order to prevent over-analyzing the limited data. The absence of measurement outcomes of both patient and caregiver concerning the same patients makes it impossible to examine correlations between reporter types. Furthermore, a reporter bias may have occurred as it was for the patients and caregivers to decide who should fill out the questionnaire. On the other hand, practical and clinically meaningful information regarding preferences of patients with CL/P and their caregivers was provided.

**Clinical Interpretations and Recommendations**

At all ages, the majority of the participants completed the CLEFT-Q scales together. In younger patients, caregivers completed the scales more often alone. The opposite is observed in older patients, where the proportion of patients completing the scales alone exceeded the “caregiver” group.

This finding might indicate the added value of a parent-version of the CLEFT-Q scales, especially for younger patients. Many pediatric QoL questionnaires have such a validated parental version available.<sup>25</sup> The current results show that parental involvement is high, as the majority of the caregivers are completing the questionnaire either with their child or alone: the availability of a parental version of the CLEFT-Q could encourage their involvement further. It might be a valuable addition for the ICHOM Standard set as well, where the caregiver’s voice currently is captured by only one validated Parent Reported Outcome Measures that assesses speech function.<sup>10</sup> Furthermore, a parental version would enable assessments earlier than 8 years as well, which is the current minimum age for patients to complete the PROM. It would enhance longitudinal data collection, minimize patient-burden, and it is simply more user-friendly to let patients and parents decide who should fill out the PROM.

Besides the patient’s and parent’s preferences to complete the CLEFT-Q together, parental involvement offers advantages from a clinical perspective as well. Wong Riff et al reported a small number of participants in their study that felt worse about how they look after filling out the appearance scales of the CLEFT-Q.<sup>30</sup> Parental involvement could provide emotional support to pediatric patients that experience

emotional impact due to their disease or disability. Moreover, the authors of the current study noticed in their outpatient clinic that patients with CL/P who filled out the scales together with a parent, reported that the questions created a reason to openly discuss topics highlighted by the CLEFT-Q scales. When taking the discrepancies of the mean outcome scores found in the current study into consideration as well, parents might not always be aware of the emotional impact of the cleft on their child. Both findings of the current study suggest that some form of parental involvement should be preferred when a pediatric PROM is used in a clinical setting. Therefore, clinicians should encourage to minimize parental involvement to an evaluation of the outcomes, after the patient filled out the PROM scales alone.

Concerning possible discrepancies between the reporter type outcomes in 12-year-olds and 15-year-olds, the “patient” group tend to score higher (thus better) on the scales in the psychosocial domain and the appearance domain. However, no significant differences between any of the reporter types were found in the current study, despite a difference in mean outcome score of almost 20 between “patient” and “caregiver” in the Speech function scale at 12 years. The lack of significance might be caused by the wide spread of outcome scores in the groups.

Discrepancies between reporter types have been examined in studies with other pediatric QoL questionnaires as well, and large systematic reviews regarding more generic pediatric QoL measurement instruments found that differences in outcomes between patient and proxy reported questionnaires are often larger in social and emotional domains, whereas functional domains seem to show more concordance.<sup>16,31,32</sup> Volpicelli et al, who examined discrepancies between patient- and parent reports regarding psychosocial function in pediatric patients with craniofacial anomalies throughout other outcome measurement instruments than the CLEFT-Q, included patients at similar age groups. In their findings, parents tend to score more negative on the psychosocial domain than children in older age-groups (12 and 15 years).<sup>33</sup> This tendency is in line with our findings, as most psychosocial scale outcomes were higher in the “patient” group than in the “caregiver” group at age 12 and 15. However, caution should be taken with the current study results, as participants were able to decide by themselves who should fill out the PROM: It could be that patients who did not experience many problems were more prone to complete the CLEFT-Q scales by themselves, resulting in a reporter group with more optimistic outcomes.

**Suggestions for future studies.** Possibilities for further, more in-depth research concerning this topic could concern examinations of discrepancies based on the gender of the parent or the child, as was done by Ooi et al in a study regarding obese children.<sup>17</sup> Moreover, research concerning the reporter group that complete the questionnaire together should be

conducted. Although studies often focus on outcome differences between patient reports and parent or proxy reports,<sup>15-17</sup> the current study points out the importance to examine this third, large group. It is difficult to determine the exact share of the patient or the parent when the scales are filled out together, which is reflected in the results of this study: outcomes from the together reporter group are not simply the mean outcomes of the patient alone and parent alone group. The presence of any response sets within the response groups, and the differences in response sets might play a key role in differences in outcomes, as children might answer differently in order to please their parents or to appear competent.<sup>34,35</sup>

Although the diversity of this reporter group might cause complications in further psychometric assessment, the current study shows that in the clinical daily practice, the large majority of the patients fill out the questionnaire together with at least one caregiver. Therefore, it is recommended for centres that have implemented a pediatric PROM to record the reporter type.

Furthermore, future studies with a prospective set-up could elucidate possible discrepancies between the outcomes of patient and caregivers, like the study of Volpicelli et al.<sup>33</sup> It would enable the collection of outcomes of both patient and parents concerning the same patients. However, a validated parental version should be available first. In case a validated parental version would be developed, it is of great importance to first examine outcomes in a set-up as mentioned above: like in other pediatric PROMs, (validated) outcomes of parent versions and patient versions have to be compared to check if they are interchangeable, or if they are only valid for comparison of outcomes within their own responder group.

In a more clinical perspective, future studies that would focus more in-depth on the implementation of both patient and (future) parent versions of the CLEFT-Q scales could be of great value for clinical use. Implementing PROM scales in an efficient and meaningful manner, even without the current presence of a parent version, is important to keep patient burden minimal while still retrieving all information that is important. Implementing the right CLEFT-Q scales at the optimal time-points might improve patient care, without “asking too much” of our patients and families, and several studies have already made recommendations to do so.<sup>36,37</sup> A critical look towards the establishment of sets of outcome measures should be maintained and, depending on the objective of the set of outcome measures, potential availability of a parent version could help in deciding whether some scales could be better completed by parents or by patients.

## Conclusion

Although the CLEFT-Q questionnaire is designed and validated for patient reports, the majority of patients with CL/P

and parents prefers to fill out the CLEFT-Q scales together. More patients complete the questionnaire alone in older age-groups, whereas the group of parents that complete the questionnaire for their child decreases with the age of the patient. Mean outcome scores per age suggest discrepancies between reporter types, in which patients tend to score more positively when completing the questionnaire alone. Therefore, caution should be taken concerning reporter type, and it is recommended to record reporter type in pediatric PROMs, in order to promote standardization of outcome measures and optimize research possibilities. From a clinical perspective, parental involvement should be encouraged in the assessment of pediatric PROMs, as involvement can motivate families to openly discuss the impact on the QoL of the child. Therefore, a recommended alternative to completing the questionnaires together, is to let parents evaluate outcomes of the PROM with their child after the PROM is filled out by the patient alone. This way, parental support can be provided, while the outcomes should not be affected by parental opinions. In the future, a parental version of the CLEFT-Q could be of additional value, especially in younger patients.

## Appendix

### Appendix 1. Kruskal Wallis H Test Results for all CLEFT-Q Scales.

CLEFT-Q scale	P-value		
	8y	12y	15y
Face	.597	.442	.257
Teeth	.858	.750	.907
Jaw	-	.438	.207
Psychological function	-	.438	.627
Social function	.183	-	-
School function	-	.106	.242
Speech function	-	.070	.867
Speech distress	-	.087	.996

### Appendix 2. Confidence Intervals of the Trend Proportions Per Age Group, Per Responder Type.

Age	Responder group	Lower bound	Upper bound
8y	Patient	0.023	0.08
	Caregiver	0.16	0.26
	Together	0.69	0.80
12y	Patient	0.05	0.13
	Caregiver	0.06	0.13
	Together	0.77	0.87
15y	Patient	0.13	0.36
	Caregiver	0.02	0.17
	Together	0.56	0.81

## Acknowledgments

We would like to thank all children and caregivers that participated in the study, that contributed by filling out the questionnaires and allowing the research team to use their data for the current study. Furthermore, we would like to thank the cleft team clinicians from both participating centers, that contributed tremendously by collecting the data consequently and by advising the research team on important clinical considerations

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## Ethical Approval

Institutional Review Board approval was obtained MEC-2016-156.

## ORCID iDs

Saranda Ombashi  <https://orcid.org/0000-0001-7988-052X>

Esther E.B. vanOers-Hazelzet  <https://orcid.org/0009-0006-8393-7748>

## Data Availability

Use of the CLEFT-Q™ questionnaire, developed by Anne Klassen, PhD, and Karen Wong, MD, MS, was made under license from McMaster University, Hamilton, Ontario, Canada. The ICHOM Standard Set for Cleft Lip/Palate is maintained under the auspices of the International Consortium for Health-Outcome Measurement (ICHOM) and is available free at <http://www.ichom.org>.

## References

- Porter M, Larsson S and Lee T. Standardizing Patient Outcomes Measurement. *New Engl J Med*. 2016;374(6):504-506.
- Squitieri L, Bozic K and Pusic A. The role of patient-reported outcome measures in value-based payment reform. *Value Health*. 2017;20(6):834-836.
- Pusic A, Lemaine V, Klassen A, Scott A and Cano S. Patient-reported outcome measures in plastic surgery: use and interpretation in evidence-based medicine. *Plast Reconstr Surg*. 2011;127(3):1361-1367.
- Slade A, Isa F, Kyte D, et al. Patient reported outcome measures in rare diseases: a narrative review. *Orphanet J Rare Dis*. 2018;13(1):61.
- Semb G, Enemark H, Friede H, et al. A scandecleft randomised trials of primary surgery for unilateral cleft lip and palate: 1. *Plan Manag J Plast Surg Hand Surg*. 2017;51(1):2-13.
- Shaw W, Semb G, Nelson P, et al. The Eurocleft project 1996-2000: overview. *J Craniomaxillofac Surg*. 2001;29(3):131-140.
- Long R, Jr, Hathaway R, Daskalogiannakis J, et al. The americleft study: an inter-center study of treatment outcomes for patients with unilateral cleft lip and palate part 1. Principles



- and study design. *Cleft Palate Craniofac J.* 2011;48(3):239-243.
8. Tucker F, Morris J, Committee J, et al. EUROCAT: an update on its functions and activities. *J Community Genet.* 2018;9(4):407-410.
  9. European Reference Network for craniofacial anomalies ERN Cranio. 2020. <https://ern-cranio.eu/>. Accessed May, 2023.
  10. Allori AC, Kelley T, Meara JG, et al. A standard set of outcome measures for the comprehensive appraisal of Cleft Care. *Cleft Palate Craniofac J.* 2017;54(5):540-554.
  11. Wong Riff KW, Tsangaris E, Goodacre T, et al. International multiphase mixed methods study protocol to develop a cross-cultural patient-reported outcome instrument for children and young adults with cleft lip and/or palate (CLEFT-Q). *BMJ Open.* 2017;7: e015467.
  12. Wong Riff K W Y, Tsangaris E, Goodacre T E E, et al. What matters to patients with cleft lip and/or palate: an international qualitative study informing the development of the CLEFT-Q. *Cleft Palate Craniofac J.* 2018;55(3):442-450.
  13. Klassen AF, Riff K W W, Longmire NM, et al. Psychometric findings and normative values for the CLEFT-Q based on 2434 children and young adult patients with cleft lip and/or palate from 12 countries. *CMAJ.* 2018;190(15):E455-E62.
  14. Tsangaris E, Wong Riff K W Y, Goodacre T, et al. Establishing content validity of the CLEFT-Q: A new patient-reported outcome instrument for cleft Lip/Palate. *Plast Reconstr Surg Glob Open.* 2017;5(4): e1305.
  15. Upton P, Lawford J and Eiser C. Parent-child agreement across child health-related quality of life instruments: a review of the literature. *Qual Life Res.* 2008;17(6):895-913.
  16. Eiser C and Morse R. Can parents rate their child's health-related quality of life? Results of a systematic review. *Qual Life Res.* 2001;10(4):347-357.
  17. Ooi DSQ, Loke KY, Ho CWL, et al. Self and parent-proxy rated health-related quality of life (HRQoL) in youth with obesity: are parents good surrogates? *Qual Life Res.* 2020;29:2171-2181.
  18. Landgraf JM and Abetz LN. Functional status and well-being of children representing three cultural groups: initial self-reports using the chq-cf87. *Psychol. Health.* 1997;12(6):839-854.
  19. Landgraf JM, Maunsell E, Speechley KN, et al. Canadian-French, German and UK versions of the Child Health Questionnaire: methodology and preliminary item scaling results. *Qual Life Res.* 1998;7(5):433-445.
  20. Ravens-Sieberer U. Der kindl-r fragebogen zur erfassung der gesundheitsbezogenen lebensqualität bei kindern und jugendlichen-revidierte form. In: *Diagnostische Verfahren zu Lebensqualität und Wohlbefinden*; (Vol. 2, pp. 184-188). Göttingen: Hogrefe.
  21. Ravens-Sieberer U, Auquier P, Erhart M, et al. European KIDSCREEN Group. The KIDSCREEN-27 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. *Qual Life Res.* 2007;16(8):1347-1356.
  22. Ravens-Sieberer U, Erhart M, Rajmil L, et al. The European KIDSCREEN Group. Reliability, construct and criterion validity of the KIDSCREEN-10 score: a short measure for children and adolescents' well-being and health-related quality of life. *Qual Life Res.* 2010;19(10):1487-1500.
  23. Ravens-Sieberer U, Morfeld M, Stein RE, et al. [The testing and validation of the German version of the impact on family scale in families with children with disabilities] Der Familien-Belastungs-Fragebogen (FaBel-Fragebogen) - Testung und Validierung der deutschen Version der "Impact on Family Scale" bei Familien mit behinderten Kindern. *Psychother Psychosom Med Psychol.* 2001;51(9-10):384-393.
  24. Graham P, Stevenson J and Flynn D. A new measure of health-related quality of life for children: Preliminary findings. *Psychol Health.* 1997;12(5):655-665.
  25. Varni JW, Seid M and Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. *Med Care.* 1999;37(2):126-139.
  26. Tsangaris E, Riff K W Y W, Vargas F, et al. Translation and cultural adaptation of the CLEFT-Q for use in Colombia, Chile, and Spain. *Health Qual Life Outcomes.* 2017;15(1):228.
  27. Klassen AF, Tsangaris E, Forrest CR, et al. Quality of life of children treated for cleft lip and/or palate: a systematic review. *J Plast Reconstr Aesthet Surg.* 2012;65(5):547-557.
  28. Hair J, Hult GTM, Ringle C and Sarstedt M. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. 2022. Sage Publishing.
  29. Mossey PA and Modell B. Epidemiology of oral clefts 2012: an international perspective. *Front Oral Biol.* 2012;16:1-18.
  30. Klassen AF, Dalton L, Goodacre TEE, et al. Impact of completing CLEFT-Q scales that ask about appearance on children and Young Adults: an International Study. *Cleft Palate Craniofac J.* 2020;57:840-848.
  31. Eiser C and Morse R. Quality-of-life measures in chronic diseases of childhood. *Health Technol Assess.* 2001;5(4):1-157.
  32. Morris C, Janssens A, Shilling V, et al. Meaningful health outcomes for paediatric neurodisability: Stakeholder prioritisation and appropriateness of patient reported outcome measures. *Health Qual Life Outcomes.* 2015;13:87.
  33. Volpicelli EJ, Pfaff MJ, Potemra HMK, et al. Discrepancies in parent perceptions and patient-reported psychosocial function in children with Craniofacial Anomalies. *Plast Reconstr Surg.* 2020;145(1):175-183.
  34. Matza LS, Swensen AR, Flood EM, Secnik K and Leidy NK. Assessment of health-related quality of life in children: a review of conceptual, methodological, and regulatory issues. *Value Health.* 2004;7(1):79-92.
  35. Annett RD. Assessment of health status and quality of life outcomes for children with asthma. *J Allergy Clin Immunol.* 2001;107(5 Suppl):S473-S481.
  36. Ombashi S, Kurniawan MSIC, Koudstaal MJ, et al. Most efficient and meaningful patient reported appearance assessment in different cleft types and age-groups with Cleft-Q. *Plast Reconstr Surg.* Published online April 11, 2023. doi:10.1097/PRS.000000000010523
  37. Apon I, van Leeuwen N, Koudstaal MJ, et al. Optimizing the psychosocial function measures in the international consortium for health outcomes measurement standard set for cleft. *Plast Reconstr Surg.* 2023;151(2):274e-281e.