

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<u>http://bmjopen.bmj.com</u>).

If you have any questions on BMJ Open's open peer review process please email <u>info.bmjopen@bmj.com</u>

Combining clinical and patient reported outcome measures for standardized speech assessment in cleft patients with the ICHOM Cleft Standard Set: What is the optimal assessment of speech in cleft lip and palate patients?

Journal:	BMJ Open
Manuscript ID	bmjopen-2023-071571
Article Type:	Original research
Date Submitted by the Author:	03-Feb-2023
Complete List of Authors:	Ombashi, Saranda; Erasmus Medical Center, plastic and reconstructive surgery Kurniawan, Melissa; Erasmus Medical Center, plastic and reconstructive surgery Allori, Alexander; Duke University Hospital & children's health center, Plastic, Maxillofacial & Oral Surgery Sharif-Askary, Banafsheh; Duke University Hospital & children's health center, Plastic, Maxillofacial & Oral Surgery Rogers-Vizena, Carolyn; Boston Children's Hospital, Plastic and Oral Surgery Koudstaal, Maarten; Erasmus MC Sophia Children Hospital, Department of Oral and Maxillofacial Surgery Franken, Marie-Christine; Erasmus Medical Center, Ear, Nose and Throat, Speech Therapy Mink van der Molen, Aebele; University Medical Center Utrecht, Department of Plastic Surgery Mathijssen, Irene; Erasmus MC Sophia Children Hospital, Department of Plastic and Reconstructive Surgery Klassen, Anne; McMaster University, Pediatrics Versnel, Sarah; Erasmus MC Sophia Children Hospital, plastic and reconstructive surgery
Keywords:	ORAL & MAXILLOFACIAL SURGERY, Paediatric head & neck surgery < PAEDIATRIC SURGERY, Paediatric oral & maxillofacial surgery < PAEDIATRIC SURGERY, Paediatric plastic & reconstructive surgery < PAEDIATRIC SURGERY, PLASTIC & RECONSTRUCTIVE SURGERY, Quality of Life

SCHOLARONE[™] Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our <u>licence</u>.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which <u>Creative Commons</u> licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

terez oni

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies



Combining clinical and patient reported outcome measures for standardized speech assessment in cleft patients with the ICHOM Cleft Standard Set: What is the optimal assessment of speech in cleft lip and palate patients?

S. Ombashi^{1,2}, M. Kurniawan^{1,}, A.C. Allori³, B. Sharif-Askary³, C.R. Rogers-Vizena⁴, M.J. Koudstaal^{2, 5, 6}, M.C. Franken⁷, A.B. Mink van der Molen^{2,8}, I.M.J. Mathijssen^{1,2,6}, A.F. Klassen⁹, S.L. Versnel^{1,2,6}

¹ Department of Plastic and Reconstructive Surgery, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands

² European Reference Network for rare and/or complex craniofacial anomalies and ear, nose and throat disorders

³ Department of Plastic, Maxillofacial & Oral Surgery, Duke University Hospital & Children's Health Center, Durham, North Carolina, United States of America

⁴ Department of Plastic and Oral Surgery, Boston Children's Hospital, Boston, Massachusetts, United States of America

⁵ Department of Oral and Maxillofacial Surgery, Special Dental Care and Orthodontics, Erasmus University Medical Center, Rotterdam, The Netherlands

- ⁶ Dutch Craniofacial Center, Erasmus University Medical Center, Rotterdam, The Netherlands
- ⁷ Department of Ear, Nose and Throat, Speech Therapy, Erasmus University Medical Center, Rotterdam, The Netherlands
- ⁸ Department of Plastic and Reconstructive Surgery, Univeristy Medical Center Utrecht, Utrecht, The Netherlands
- ⁹ Department of Pediatrics, McMaster University, Hamilton, Canada
- Numer of words: 4031

- Number of tables: 2
- Number of figures: 4
- Number of appendices: 3

Corresponding Author: S. Ombashi, MD. Dr. Molewaterplein 40, 3015 GD Rotterdam, Erasmus University Medical

Center; Tel: +316-31989928; E-mail: s.ombashi@erasmusmc.nl

ABSTRACT

Objectives

Speech problems in patients with a cleft palate are often complex and multifactorial. Finding the optimal way of monitoring these problems is challenging. The ICHOM(International Consortium of Health Outcomes Measurement) developed a set of standardized outcome measures at specific ages for patients with a cleft lip and/or palate, including measures of speech assessment. This study evaluates the type and timing of speech outcome measures currently included in this ICHOM set, and additionally discusses speech assessments in other cleft protocols and initiatives.

Methods

An international, multicenter study was set up. Outcomes of clinical measures and Patient Reported Outcome Measures(PROMs) were collected according to the ICHOM set. PROM data from a field test of the CLEFT-Q were collected to examine the value of additional moments of measurement that are used in other cleft initiatives. Analyses were done per cleft type and in different age-groups, and included univariate regression analyses, trend analyses, T-tests, correlations and floor and ceiling effects.

Results

A total of 2500 patients were included. The PROMs correlated low to moderate with clinical outcome measures. The clinical outcome measures correlated low to moderate with each other as well. In contrast, two CLEFT-Q scales correlated strongly with each other. All PROMs and the Percent Consonants Correct(PCC) showed an effect of age. In patients with an isolated cleft palate, a ceiling effect was found in the intelligibility in context scale(ICS).

Conclusion

Recommendations for an optimal speech outcome assessment in cleft patients are made, based on the evaluation of the speech coutome measures in the ICHOM set . Measurement moments of different cleft protocols and initiatives are considered in this proposition. Concerning the type of measures, adjustment of the current PCC score outcome seems appropriate. For centres with adequate resources and specific interest in research, translation and validation of an upcoming tool, the CAPS-A, is recommended.

Strengths and limitations

- International multicenter study
- Cross-sectional data analyses
- Elaborate evaluation of multiple time-points
- Assessment of PROMs and clinical outcome measures

INTRODUCTION

In cleft palate patients with or without a cleft lip (CP±L), speech development is often complex. Persistent velopharyngeal incompetence, residual fistula, adenoid atrophy, surgical intervention and hearing problems influence speech disorder severity in this population(1-5). Speech problems in patients with CP±L can have a large impact on an individual's life, as proper speech skills play an essential role in activities, social functioning and participation in society(6). Many treatment pathways are focused on speech improvement to ameliorate Quality of Life (QoL)(7). Logically, speech assessment is an important parameter in cleft care.

However, no consensus has been reached regarding best diagnostic speech outcome measures and their timing in this population(5). Developing scientifically solid instruments to assess speech in an objective manner is complicated, because listener's perception of speech deficits, even by experts, may differ substantially(8). An additional challenge is systematic assessment of the patient's perspective, which is essential to include due to the impact of speech problems on the individual(9). Although widely-accepted agreement seems essential for improvement of cleft care, finding consensus is complex, especially since speech outcomes should be comparable between different languages to facilitate international collaboration.

Recently, the International Consortium for Health Outcomes Measurement (ICHOM) developed the ICHOM Standard Set for Cleft Lip and Palate (ICHOM Standard Set), with different pathways for varying cleft types(10). Based on patient and expert consensus, a minimal, accessible set of outcome measures was established to enable benchmarking between cleft centers in a systematic manner. For speech assessment, an outcome set was included with both clinical measures and Patient Reported Outcome Measure (PROMs), being the patient's and parent's perspectives.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

BMJ Open

> So far, the selected standardized speech outcome measures and their timing have not been evaluated. As an increasing number of centers are implementing this set, it is important to critically evaluate and optimize this ICHOM Standard Set. Three centers, the Boston Children's Hospital (Boston, USA), Duke University Hospital (Durham, USA), and the Erasmus Medical Center (Rotterdam, The Netherlands), started clinical implementation and an international collaboration in 2015. The overarching aim of this collaboration is to share data and knowledge obtained by using the set in standard care. Additionally, they collaborate with McMaster University (Hamilton, Canada), who developed and field tested the CLEFT-Q questionnaire, of which many scales are included in the ICHOM Standard set.

> The objective of this study was to evaluate the current standardized speech outcome measures of the ICHOM Standard Set for patients with CP±L. More specifically, the value of every speech outcome measure was examined, as well as the best age intervals for assessment of these outcome measures. In addition, other speech assessment tools are discussed. Finally, recommendations are made for an optimal and complete assessment of speech in patients with CP±L, that is efficient and accessible for all cleft centers.

METHODS

Patient population

Three centers (Boston Children's Hospital, Duke University Hospital, and Erasmus Medical Center,) each implemented the ICHOM Standard Set in 2015. All patients treated at these centers for a cleft palate with a cleft lip/cleft alveolus (CL(A)P), or an isolated cleft palate (CP) who were assessed according to the ICHOM Standard Set (age range 5-22 years), were included. In addition, another patient group derived from an international field test of the CLEFT-Q, by McMaster University(11). According to the age cut-off of the ICHOM Standard Set, only outcomes from CLEFT-Q scales of field test patients with a CP±L up to 22 years old were included in the current study(fig 1). Patients from the participating centers were excluded in case they could not sufficiently speak or write the language native to the center's é. C.C. country.

Patient and public involvement

In the development of the ICHOM Standard Set, patients were actively involved. The ICHOM Standard Set was implemented in each center as part of regular clinical care. Data was pseudonymized and collected retrospectively, and ethical approval was obtained to do so without explicit consent from each patient and parent. Results of this study may be of use to further improve the currently used ICHOM Standard Set, and therewith regular clinical care.

Outcome measures

PROMs

CLEFT-Q Scales

The CLEFT-Q is developed specifically to assess QoL from the patient's perspective in patients with a CP±L. A literature review, patient interviews and psychometric testing, established the final content of the scales, which covers several overarching domains(12-14). Speech is assessed through two scales, each covering a different domain. Both scales have 3 response options for each item (always; sometimes; never); a lower score equals a worse outcome. Completing the scales can be done online; it will take the patient several minutes.

Speaking-Related Distress (SDistress) is part of the psychosocial domain. The scale contains 10 items that relate to the psychosocial part of speech difficulties, like nervousness or frustration(14). Speech Function (SFunction) focuses on the functional speech difficulties that patients themselves identify, for example the ability to say certain letters or words. The scale consists of 12 items that belong to the facial function domain(14).

<u>Intelligibility in Context Scale(ICS)</u> is a 7-item, parent-reported questionnaire to assess the intelligibility of the child on a Likert-point scale. The items assess the degree to which the speech of the patient is understood by different communication partners. ICS appeared to be a valid and reliable tool for children with speech disorders, (15, 16), but not specifically designed nor validated for patients with CP±L(17). It is available in several languages, and normative data exists for English-speakers(18, 19).

Clinical outcome measures

<u>Percent Consonants Correct(PCC)</u> is developed to detect speech sound errors. PCC scores are calculated by using a standard, crossecitonally translated set of words that include all speech problems children with CP±L often tend to have.

BMJ Open

In case of any problems, their severity can be categorized: PCC scores of 85-100% indicate mild to no problems; scores of 65-84.9% indicate mild-moderate problems; scores of 50-64.9% indicate moderate-severe problems; and scores <50% indicate severe problems(20). PCC is suitable for usage in patients with CP±L when assessed by well-trained clinicians(8).

Velopharyngeal Competence rating(VPC) discriminates between three categories: 'competent',

'marginally incompetent' and 'incompetent'. The outcome is determined by the speech therapist based on the PCC test and spontaneous speech. In case of any clinical evidence of minor problems regarding the competence, VPC was categorized as 'marginally incompetent'. When clinically significant problems were detected, suggesting surgical management and/or speech therapy, VPC was categorized as 'incompetent'. Prior studies found VPC to be suitable as a first clinical choice for the assessment of velopharyngeal dysfunction and is recommended for both clinical follow-up and research(20).

Data collection

All participating centers obtained ethical approval for the current study from their local ethics committees. Data was collected restrospectively over a 6 year period (2015-2020). According to the ICHOM protocol, both CLEFT-Q Speech scales were assessed at ages 12 and 22 years(fig. 1). Both PCC and VPC were scored at ages 5, 12 and 22 years, and ICS at ages 5 and 12 years. The field test cross-sectionally collected data from patients with a cleft across 12 different countries with different income-statuses(11). As 8 years is the minimum age to complete the CLEFT-Q, both CLEFT-Q Speech scales from field test patients with CP±L from 8-22 years old were included(fig. 1). Income status of the country according to the World Bank Classification was made within the field test. Data from the ICHOM centers were all categorized as deriving from high-income countries.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Baseline characteristics that were collected included gender, type of cleft, and age at the time of assessment.

Data analysis

Data was analysed in R-studio, a free software environment for statistical computing and graphics(21). Psychometric validation of the SDistress and SFunction confirmed suitability to use a 0 to 100 scale deriving from the sum scores for analysis(11).

For analysis of the ICS questionnaire, the average score of the seven items was used. VPC was used as an ordinal variable, whereas PCC scores were expressed as proportions.

All participating ICHOM centers are high-income countries, whereas part of the field test data was collected in upper middle and lower middle income countries. To prevent possible influence of incomestatus on the outcomes, univariate regression analyses were used to examine differences in outcome scores of the SDistress and SFunction before further analyses. Data was categorized according to the income status of the country where the data had been collected.

In order to examine the added value of each PROM and clinical outcome measure in regard with the other measures, correlations were examined between the PROMs; between the clinical outcome measures; and between the PROMs and the clinical outcome measures. Pearson correlations were used, and outcomes were analysed per cleft type. Correlations were considered strong in case r>0.7; moderate between r=0.5 and R=0.7; and weak in case r<0.5.

Analyses within and between different age-groups were done to explore whether the current outcome

measures are assessed at the optimal age-points, and whether additional measurement moments are indicated either for PROMs or clinical outcome measures. Therefore, not only time-points of the ICHOM protocol were included in analyses. As CLEFT-Q outcome scores of all ages between 8-22 years were included from the field test data, time-points used by other large initiatives as Eurocleft, Scandcleft and

Americleft were considered as well (22-24). Doing so, the following age-groups were set up: 5-7 years; 8-9 years; 10-13 years; 14-16 years; 17-19 years, 20-22 years(fig. 1).

Per age-group, possible differences between scale scores were examined with independent T-tests. Bonferroni correction was applied for multiple testing. Trend analyses were performed to identify potential problems in specific age-groups.

Floor- and ceiling effects were examined to identify the suitability of the outcome measures in our population. A floor or ceiling effect is seen when a considerable amount of the outcome scores are either scored the best (in this case a maximum score, thus a ceiling effect), or the worse (in this case a minimum score, thus a floor effect). Both effects result in a truncated distribution of the outcomes on either side of the scale (25, 26). Minimum and maximum score outcomes of all PROMs and the PCC were evaluated. A percentage of 20% or more of the patients scoring the minimum or maximum outcome score was considered as a ceiling effect. In VPC, the outcome distributions were examined. BMJ Open: first published as 10.1136/bmjopen-2023-071571 on 28 December 2023. Downloaded from http://bmjopen.bmj.com/ on June 14, 2025 at Agence Bibliographique de Enseignement Superieur (ABES)

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

Characteristics of the included participants

A total of 2500 patients were included in the study; 1723 derived from the field test, and 777 from the

ICHOM centers (table 1).

Table 1: Demographics and phenotypes

Descriptives total

	Field test (n = 1723)	ICHOM (n = 777)	Overall (n = 2500)
Sex			
Male	981 (56.9%)	444 (57.1%)	1425 (57.0%)
Female	742 (43.1)	333 (42.9%)	1075 (43.0%)
Cleft Type			
СР	517 (30.0%)	301 (38.7%)	818 (32.7%)
CL(A)P	1206 (70.0%)	476 (61.3%)	1682 (67.3%)
Income classifications		0	
High Income	1364 (79.2%)	777 (100%)	2141 (85.6%)
Up Middle income	199 (11.5%)	0 (0%)	199 (8.0%)
Low Middle Income	160 (9.3%)	0 (0%)	160 (6.4%)

There were slightly more males than females, and relatively more patients with a CL(A)P than with a CP. Significant differences between countries with a High and Upper-/Lower-Middle income status of the field test were found and results are shown in appendix 1.

Therefore, further analyses were done only with the patient population of the field test deriving from countries with a high income status, like all participating ICHOM Cleft centers (n = 2141). The subgroup characteristics are included in appendix 2.

Associations between the	e outcome measures
--------------------------	--------------------

Correlations between all outcome measures (clinical and PROMs), in both cleft types (CP and CL(A)P), appeared significant (P<0.05), except for the correlation between PCC and SDistress (p = 0.285)(fig. 2). <u>Correlation PROMs</u>: The SDistress and SFunction showed a strong correlation in patients with CP (r = 0.76) and a moderate correlation in patients with CL(A)P (r = 0.68). The ICS and SFunction correlated strongly (r = 0.73) in patients with CP, and moderately (r = 0.64) in patients with CL(A)P; whereas the ICS and the SDistress correlated moderately in patients with CP (r = 0.52) and weak in patients with CL(A)P (r = 0.47).

<u>Correlation clinical outcomes</u>: VPC and PCC were (negatively) moderately correlated in both cleft types (r = -0.62 and r = -0.67 in patients with CP and CL(A)P respectively).

<u>Correlation PROMs and clinical outcomes</u>: Moderate correlations were found between the PCC and ICS in patients with CP (r = 0.64), and in patients with CL(A)P (r = 0.5). VPC and ICS had a (negative) weak correlation in patients with CP (r = -0.49), and CL(A)P (r = -0.43). The SDistress and SFunction were weakly correlated with VPC and PCC (negatively) in both cleft types (fig. 2).

Comparing outcome measures between age-groups

<u>SDistress and SFunction</u>: showed the highest mean outcome (i.e. the most favourable ratings) in the age-group of 14-16 years old. From thereon, a slightly downward trend is seen (fig. 3). In both CLEFT-Q Speech scales the lowest mean outcome scores were found in the age-group of 8-9 years old, which was significantly different in comparison with the other age-groups in patients with CL(A)P (p<0.05, table 2).

	Age categories	SDistress	SFunction	ICS	PCC
te te	5 – 7			4.08 (0.73)	72.49 (31.14)
Cle. ala	8-9	70.44 (20.64)	66.12 (22.12)		
- C	10 – 13	74.84 (19.91)	72.56 (21.52)	4.45 (0.62) ¹	92.02 (11.48) ¹

Table 2 Mean outcomes per age-group, per cleft type

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

	14 – 16	75.10 (22.77)	74.66 (24.32)		
	17 – 19	70.59 (16.92)	73.00 (19.02)		
	20 – 22	70.27 (21.93)	71.69 (24.98)		90.92 (15.33) ¹
	5 – 7			4.03 (0.50)	46.06 (25.59)
	8-9	65.32 (21.61)	64.75 (20.16)		
e Lip	10 - 13	73.16 (20.62) ¹	69.57 (20.96)	4.32 (0.55) ¹	76.25 (22.63) ¹
eft I alat	14 – 16	72.81 (19.66) ¹	74.22 (19.14) ¹		
B B	17 – 19	71.82 (19.67) ¹	71.41 (20.79) ¹		
	20 – 22	71.49 (18.00)	72.26 (18.40) ¹		86.83 (16.13) ^{1,2}

p < 0.05 compared to 5-7 or 8-9 age

P < 0.05 compared to 10 - 13 age

ICS : Both patient groups showed a significant difference between the two age-groups 5 and 12 years, ICS was significantly lower at 5 years than at 12 years in both cleft types (tab. 2).

PCC: Observing the trends in the clinical outcome measures, an upward trend regarding PCC score was

seen (fig. 3). In both cleft types, PCC scores differed significantly between the age-groups(tab. 2).

VPC: In the age-group of 5 years, 25.6% of the patients with CP and 60.6% of the patients with CL(A)P

were scored as incompetent. In 22 year-olds, this percentage was 11.1% in patients with CP and 16.7%

in patients with CL(A)P.

No floor effects were found in any of the PROMs. In patients with CP, the ICS showed a ceiling effect

(29.0%, n = 169). No ceiling effects were observed in patients with CL(A)P. An overview of all maximum

scores and the VPC score distribution is shown in appendix 3.

DISCUSSION

Evaluation of the value of the current ICHOM speech outcome measures

All correlations between PROMs were moderate, except for the strong correlation of the SFunction with both the SDistress and the ICS in patients with a CP. The fact that the correlation between the SFunction and SDistress is stronger in patients with CP than in patients with CL(A)P, suggests that the visibly different appearance in patients with CL(A)P plays a significant role in SDistress as well; in a social context, looking differently may cause additional or more distress besides having speech problems. This is supported by our finding that the ICS correlated moderately with SFunction, but weakly with SDistress in the CL(A)P group. Parent reported speech intelligibility correlated higher to children's self report of their speech function, than it did to the speech distress the children themselves experience. In the latter, distress about appearance could be included. This finding suggests that the ICS can give an indication of 'patient-reported' SFunction in young children who cannot complete a PROM themselves yet (7 years and younger).

The PROMs showed weak correlations with the clinical reported outcomes; except for the moderate correlation that was seen between the ICS and the PCC in both patient groups. Based on these findings, PROMs appear to be of added value, as they provide different information than captured with the clinical outcome measures included in the Standard Set. They add a unique dimension to speech outcome measurement – a subjective dimension related to the patient's experiences with everyday speaking situations. While clinical measures objectively appraise the quality of speech, they will probably be insufficient to adequately capture the more nuanced social, emotional, and psychological aspects of SDistress and SFunction. With this additional self-report and parental information, clinicians

can more comprehensively explore the patients' problems concerning speech in order to find out whether additional treatment or guidance is indicated.

Evaluation of the impact of age of assessment on measurement outcomes

In both CLEFT-Q Speech scales, the age-group of 8-9 years enholds the worst scores. Speech improvement due to speech therapy or late closure of the hard palate (in certain protocols around the age of 9 years when alveolar bone grafting is performed), might explain the higher, better scores in the age-groups of 10-13 and 14-16 years. In age-groups 17 and up however, CLEFT-Q scores appeared to decline whereas PCC scores improved. This finding suggests that (almost) adult patients with CP±L develop feelings of insecurity concerning their speech, although their speech sound production remains good, or even improves. This is in line with speech therapists' experiences in the outpatient clinic, where patients were seen in person at the age of 22, but not at age of 17-19. Quite often, when discussing outcomes of the CLEFT-Q scales as well as the PCC with the patient, (s)he reacted surprised when told that no (cleft-related) problems were present in their speech. Taking the lower CLEFT-Q scores in 8-9-, 17-19- and 20–22-year-olds that were found in the field test into consideration, additional assessment of a PROM at the age-groups of 8-9 (youngest age at which this PROM can be assessed) and 17-19 years old should be considered for implementation in the ICHOM Standard Set. Therewith, monitoring patients more closely will be enabled, and any concerns of patients with CP±L regarding their speech can be discussed timely.

The two CLEFT-Q speech scales showed to capture overlapping information as they strongly correlate in patients with CP. Questions deriving from the SDistress are not measurable in any other manner, whereas SFunction from the patient's perspective might be less of added value for a PROM questionnaire. Therefore, implementation of the CLEFT-Q SDistress scale in patients with both cleft types is recommended at the age-groups of 8-9 years and 17-19 years. (fig. 4).

BMJ Open

A ceiling effect in **ICS** outcomes of patients with CP, without clear differences between average scores in patients with CP and CL(A)P, suggests that the group with CP contains a diverse population and severity of the speech problems vary widely. Furthermore, since ICS is not specifically developed for a population with CP±L, it is debatable whether this tool captures the information necessary to point out all relevant speech problems in the patient group.

However, exclusion of ICS could mean that a large part of the speech problems in the population with CP would remain undetected. Assessment at 5 and 12 years in patients with both cleft types, which is the current timing in the ICHOM Standard Set, appears therefore appropriate despite the ceiling effect.

Although **VPC** scores were relatively favourable in patients with CP, no changes regarding the implementation of the VPC scores are recommended as the outcomes showed to vary. VPC can serve as a suitable screening tool and outcomes are easily gathered by the observation of a clinician. Hence, patient-burden is low and the tool efficiently detects any velopharyngeal problems.

PCC scores that were found indicated speech sound problems especially in the younger age-groups of the patients with CL(A)P. 22-year-olds with both cleft types showed mild speech sound problems in general. Therefore time points as currently implemented in the ICHOM Standard Set appear adequate. In contrast, the suitability of PCC-assessment in a cleft set focusing on standardized outcome measures is still debatable, as inter- and intra-center reliability have not been investigated thoroughly in all participating centers so far(8). Future research should include an examination of scoring and interpreting PCC scores in different centers and/or different countries.

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

Future considerations regarding alternative speech outcome measures

In order to establish an optimal cleft set for speech assessment, other standardized outcome measures should be considered. Based on clinical experience with ICHOM Standard Set, possible suggestions for additional outcome measures are discussed here.

Regarding **PROMs** for speech assessment in patients with CP±L, the CLEFT-Q scales seem to be the most suitable PROMs available. Their comprehensive psychometric examination and cross-cultural character make them accessible for all cleft centers that seek an efficient minimal cleft set that comprises all important speech parameters (11-13). The standardized approach for translation and validation of the CLEFT-Q questionnaire enables accessibility of the PROM even for centers that still need to translate the CLEFT-Q into their native language(27, 28). Another cleft-specific PROM is the Cleft Hearing and Speech Questionnaire (CHASQ). Whereas the psychometric properties of the CLEFT-Q were examined throughout Rasch measurement theory, classical test theory was used for the CHASQ(29). A recent cross-sectional questionnaire study that compared the CLEFT-Q with the CHASQ, found that the majority of the patients with CP±L preferred the CLEFT-Q(29). Therefore, implementation of the CHASQ speech does not seem to be of added value in the current cleft set.

Besides the used **VPC** measure, a more elaborate variant exists, namely the VPC-Summary (VPC-Sum). This includes assessment of hypernasality, passive VPI symptoms and the transcriptions of active nonoral consonant errors(30). VPC-sum can either be reported as a score between 0 and 6, or as a dichotomized outcome (velopharyngeal competence or incompetence)(30).

Calculation of the VPC-Sum is based on single-words, whereas VPC-rate is based on observation of spontaneous speech (31). VPC-Sum would be an interesting measure due to its efficiency, although it may not be achievable to implement VPC-Sum in all centers in the near future as only five different

BMJ Open

languages are currently available(31). Other alternatives such as nasopharyngoscopy or MRI are invasive, expensive and enlarge the patient burden(32), and therefore not easy accessibility for all centers.

The currently implemented **PCC** lacks any categorization of consonant errors. The **Eurocleft speech group** created a research protocol with a phonetic framework, which was used in six centers and five different languages(33). It also included consonant production, but assessed on sentence level instead of single words. It categorized into 3 groups (correct; almost correct; incorrect). Further division into 21 error categories that were sampled in five groups was done in case of incorrect consonants (Nasal airflow; glottal realisations; alveolar deviations; sibilant deviations; other)(33). Moreover, general speech quality was assessed concerning hyper- and hyponasality, and voice quality(34). Expert rating of these outcomes requires periodic training of sufficient interrater reliability. However, it might be too detailed for implementation in an efficient, clinically oriented cleft set. Therefore, we suggest to further categorize the PCC score, although not as detailed as in the Eurocleft studies. Based on clinical experience with the ICHOM Standard Set, it is recommended that speech pathologists report whether any cleft related, phonological, or phonetic problems are detected.

Another clinical outcome measure, the **Great Ormond Street Speech Assessment 1998** (GOS.SP.ASS'98), provides a comprehensive view of all speech associated features for patients with CP±L(35, 36). Its suitability for inter-center comparison would make it interesting for the ICHOM Standard Set(10); however, it is too detailed for clinical audit(37). In succession the **Cleft Audit Protocol for Speech Augmented** (CAPS-A) was developed for cleft-related problems, and could be an alternative for PCC(38). Seen its rigorous psychometric assessment, it fits well into a set that seeks standardized outcome measures. The Americleft Speech Project found that an acceptable inter- and intrarater reliability can be

Protected by copyright, including for uses related to text and

data mining, Al training, and similar technologies

BMJ Open

> achieved(37, 39). Furthermore, it is suitable for assessment in 5-years-olds, which enables detection of speech problems at an earlier age(40). However, the CAPS-A is limited in types of statistical analyses due to the scaling type used (equal appearing interval)(41). A more practical challenge concerning implementing the CAPS-A would be the required training of all involved speech therapists, and the amount of time the assessment takes (15 min.)(38). Moreover, the CAPS-A is developed and applicable for English-speaking countries, necessitating translation and validation in other languages(37). The CAPS-A is not ideal for centers interested in a minimal and efficient cleft set. However, centers with experience and resources are highly recommended to implement this tool in order to promote further international standardization of elaborate speech assessment in patients with CP±L(fig. 4). Implementation of the CAPS-A would also enable the use of the recently developed and validated CAPS-A-VPC-Sum score to reliably measure velopharyngeal function(42). Our suggestion for centers that consider the implementation of the CAPS-A, is to assess it at the ages of 5-7, 10-13 and 20-22 years in order to enable long-term follow-up.

Limitations of the study

Data was analyzed cross-sectional. Longitudinal analyses to explore development of speech and for benchmarking will be possible in the future, since data collection continues.

CONCLUSION

From the current study, it can be concluded that the current ICHOM Standard Set is informative and efficient. PROMs were shown to be of added value, and the CLEFT-Q appeared to be the most suitable PROM. Therefore, continuation of collecting the current outcome measures and time points is

BMJ Open

recommended. Furthermore, a minor extension is suggested: In addition to the current time-points of assessment, it is recommended to implement the CLEFT-Q SDistress scale at the age of 8-9 and 17-19 as well. Further adjustments of the set could comprise an additional categorization of the PCC score, based on the framework of Eurocleft and adjusted for clinical usage.

ACKNOWLEDGMENTS

We would like to thank all children and caregives 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 speech pathologists from all our participating centers, that contributed tremendously by collecting the data consequently and by advising the research team on important clinical considerations

Contributor statement Saranda **Ombashi**: Data collection in Dutch Center; correspondence to other centers; data analyses; writing and adjusted the manuscript based on advice of the other authors.* Melissa S.I.C. Kurniawan: Data analyses, co-writer on the methods and results section of the manuscript.* Alexander C. Allori: Data collection at Duke University Hospital, advisor of project plan and manuscript.* B. Sharif-Askary: Data collection at Duke University Hospital, advisor of project plan and manuscript.* Carolyn R. Rogers: Data collection at Harvard University Hospital, advisor of project plan and manuscript.* Maarten J. Koudstaal: Writer of the Ethical Board Approval, Supervision of data analyses. Advisor and supervisor on project plan and manuscript.* M.C. Franken: Data collection at Erasmus Medical Cewnter, advisor of project plan and manuscript, with special focus on the clinical content.* A.B. Mink van der Molen: advisor of project plan and manuscript.* Irene M.J. Mathijssen: Overall supervision of project plan and manuscript.* Anne F. Klassen: Data collection Field Test Mc Master University, advisor on project plan and manuscript.* Sarah L. Versnel: Project leader, correspondence with all other centers. Supervision on project plan, data analyses and manuscript.* *All authors approved the final version to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Competing interests:

The CLEFT-Q is owned by McMaster University and the Hospital for Sick Children. Prof. Anne Klassen is a co-developer of the CLEFT-Q and would receive a share of any licence revenues as royalties based on her institutions' inventor sharing policy if the CLEFT-Q was used in a for profit study. The other authors have no disclosures.

Funding

This project received no funding. However, S. Ombashi works as a PhD-student for an European Union-funded Network, the 'Eureopean Reference Network for Craniofacial anomalies and Ear, Nose and Throat disorders'. The results and findings from this project are of help in further European alignment concerning standardized outcome measures in cleft care.

Data sharing statement

A data sharing agreement was set up and signed by all participating centers that shared data for this project.

Ethics approval statement

All participating centers obtained ethical approval for the current study from their local ethics committees.

References

1. Kummer AW. Types and causes of velopharyngeal dysfunction. Semin Speech Lang. 2011;32(2):150-8.

2. Knight J, Cassell CH, Meyer RE, Strauss RP. Academic outcomes of children with isolated orofacial clefts compared with children without a major birth defect. Cleft Palate Craniofac J. 2015;52(3):259-68.

3. Van Lierde KM, Dhaeseleer E, Luyten A, Van De Woestijne K, Vermeersch H, Roche N. Parent and child ratings of satisfaction with speech and facial appearance in Flemish prepubescent boys and girls with unilateral cleft lip and palate. Int J Oral Maxillofac Surg. 2012;41(2):192-9.

4. Peterson-Falzone SJ. Speech outcomes in adolescents with cleft lip and palate. Cleft Palate Craniofac J. 1995;32(2):125-8.

5. Baek RM, Kim BK, Jeong JH, Ahn T, Park M, Han J. The effect of age at surgery and compensatory articulation on speech outcome in submucous cleft palate patients treated with double-opposing Z-plasty: A 10-year experience. J Plast Reconstr Aesthet Surg. 2017;70(5):646-52.

6. Markham C, van Laar D, Gibbard D, Dean T. Children with speech, language and communication needs: their perceptions of their quality of life. Int J Lang Commun Disord. 2009;44(5):748-68.

7. Nicola K, Watter P. Health-related quality of life from the perspective of children with severe specific language impairment. Health Qual Life Outcomes. 2015;13:127-.

8. Sell D, Sweeney T. Percent Consonant Correct as an Outcome Measure for Cleft Speech in an Intervention Study. Folia Phoniatr Logop. 2019:1-9.

9. Klassen AF, Tsangaris E, Forrest CR, Wong KW, Pusic AL, Cano SJ, 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):547-57.

10. Allori AC, Kelley T, Meara JG, Albert A, Bonanthaya K, Chapman K, et al. A Standard Set of Outcome Measures for the Comprehensive Appraisal of Cleft Care. Cleft Palate Craniofac J. 2017;54(5):540-54.

11. Klassen AF, Riff KWW, Longmire NM, Albert A, Allen GC, Aydin MA, 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.

12. Wong Riff KW, Tsangaris E, Goodacre T, Forrest CR, Pusic AL, Cano SJ, 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(1):e015467.

13. Tsangaris E, Wong Riff KWY, Goodacre T, Forrest CR, Dreise M, Sykes J, 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.

14. Wong Riff KWY, Tsangaris E, Goodacre TEE, Forrest CR, Lawson J, Pusic AL, 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):442-50.

4

5

6 7

8

9

10

11

12 13

14

15

16

17

18 19

20

21

22

23 24

25

26

27

28

29

30

31

32

33

34 35

36

37

38

39

40 41

42

43

44

45 46

47

48

49

50

51 52

53

60

BMJ Open

15. McLeod S, Crowe K, Shahaeian A. Intelligibility in Context Scale: Normative and Validation Data for English-Speaking Preschoolers. Lang Speech Hear Serv Sch. 2015;46(3):266-76. 16. McLeod S. Intelligibility in Context Scale: cross-linguistic use, validity, and reliability. Speech, Language and Hearing. 2020;23(1):9-16. 17. McLeod S, Harrison LJ, McCormack J. The intelligibility in Context Scale: validity and reliability of a subjective rating measure. J Speech Lang Hear Res. 2012;55(2):648-56. McLeod S, Harrison LJ, McCormack J. Skattning av förståelighet i kontext: Svenska 18. [Intelligibility in Context Scale: Swedish] (T. Lagerberg, Trans.) Bathurst, NSW, Australia: Charles Sturt University. 2012. 19. McLeod S, Harrison LJ, McCormack J. Schaal voor verstaanbaarheid in de context: Nederlands [Intelligibility in Context Scale: Dutch] (A. van Doornik, Trans). Bathurst, NSW, Australia: Charles Sturt University. 2013. 20. Lohmander A, Hagberg E, Persson C, Willadsen E, Lundeborg I, Davies J, et al. Validity of auditory perceptual assessment of velopharyngeal function and dysfunction – the VPC-Sum and the VPC-Rate. Clinical Linguistics & Phonetics. 2017;31(7-9):589-97. 21. The R Project for Statistical Computing: The R Foundation; 2021 [Available from: https://www.r-project.org/. Semb G, Enemark H, Friede H, Paulin G, Lilja J, Rautio J, et al. A Scandcleft randomised 22. trials of primary surgery for unilateral cleft lip and palate: 1. Planning and management. J Plast Surg Hand Surg. 2017;51(1):2-13. Long RE, Jr., Hathaway R, Daskalogiannakis J, Mercado A, Russell K, Cohen M, et al. The 23. 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):239-43. 24. Shaw WC, Semb G, Nelson P, Brattstrom V, Molsted K, Prahl-Andersen B, et al. The Eurocleft project 1996-2000: overview. J Craniomaxillofac Surg. 2001;29(3):131-40; discussion 41-2. 25. Lim CR, Harris K, Dawson J, Beard DJ, Fitzpatrick R, Price AJ. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. BMJ Open. 2015;5(7):e007765-e. Stucki G, Liang MH, Stucki S, Katz JN, Lew RA. Application of statistical graphics to 26. facilitate selection of health status measures for clinical practice and evaluative research. Clin Rheumatol. 1999;18(2):101-5. 27. Tsangaris E, Riff K, Vargas F, Aguilera MP, Alarcon MM, Cazalla AA, 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. 28. Tsangaris E, Wong Riff KWY, Dreise M, Stiernman M, Kaur MN, Piplani B, et al. Translation and cultural adaptation of the CLEFT-Q into Arabic, Dutch, Hindi, Swedish, and Turkish. European Journal of Plastic Surgery. 2018;41(6):643-52. Stiernman M, Klintö K, Persson M, Becker M. Comparison of Corresponding Scores From 29. the Cleft Hearing Appearance and Speech Questionnaire (CHASQ) and CLEFT-Q in Swedish Patients With Cleft Lip and/or Palate. The Cleft Palate-Craniofacial Journal. 2020:105566562096412.

30. Hammarström IL, Nyberg J, Alaluusua S, Rautio J, Neovius E, Berggren A, et al. Scandcleft Project Trial 2-Comparison of Speech Outcome in 1- and 2-Stage Palatal Closure in 5-Year-Olds With UCLP. Cleft Palate Craniofac J. 2020;57(4):458-69.

31. Lohmander A, Willadsen E, Persson C, Henningsson G, Bowden M, Hutters B. Methodology for speech assessment in the Scandcleft project--an international randomized clinical trial on palatal surgery: experiences from a pilot study. Cleft Palate Craniofac J. 2009;46(4):347-62.

32. Shprintzen RJ, Marrinan E. Velopharyngeal insufficiency: diagnosis and management. Curr Opin Otolaryngol Head Neck Surg. 2009;17(4):302-7.

33. Brondsted K, Grunwell P, Henningsson G, Jansonius K, Karling J, Meijer M, et al. A phonetic framework for the cross-linguistic analysis of cleft palate speech. Clinical Linguistics & Phonetics. 1994;8(2):109-25.

34. Grunwell P, Brondsted K, Henningsson G, Jansonius K, Karling J, Meijer M, et al. A sixcentre international study of the outcome of treatment in patients with clefts of the lip and palate: the results of a cross-linguistic investigation of cleft palate speech. Scand J Plast Reconstr Surg Hand Surg. 2000;34(3):219-29.

35. Sell D, Harding A, Grunwell P. A screening assessment of cleft palate speech (Great Ormond Street Speech Assessment). Eur J Disord Commun. 1994;29(1):1-15.

36. Sell D, Harding A, Grunwell P. GOS.SP.ASS.'98: an assessment for speech disorders associated with cleft palate and/or velopharyngeal dysfunction (revised). Int J Lang Commun Disord. 1999;34(1):17-33.

37. John A, Sell D, Sweeney T, Harding-Bell A, Williams A. The cleft audit protocol for speech-augmented: A validated and reliable measure for auditing cleft speech. The Cleft palate-craniofacial journal : official publication of the American Cleft Palate-Craniofacial Association. 2006;43(3):272-88.

38. Ahl R, Harding-Bell A. Comparing Methodologies in a Series of Speech Outcome Studies: Challenges and Lessons Learned. The Cleft Palate-Craniofacial Journal. 2017;55(1):35-44.

39. Chapman KL, Baylis A, Trost-Cardamone J, Cordero KN, Dixon A, Dobbelsteyn C, et al. The Americleft Speech Project: A Training and Reliability Study. Cleft Palate Craniofac J. 2016;53(1):93-108.

40. Britton L, Albery L, Bowden M, Harding-Bell A, Phippen G, Sell D. A cross-sectional cohort study of speech in five-year-olds with cleft palate ± lip to support development of national audit standards: benchmarking speech standards in the United Kingdom. Cleft Palate Craniofac J. 2014;51(4):431-51.

41. Baylis A, Chapman K, Whitehill TL, The Americleft Speech G. Validity and Reliability of Visual Analog Scaling for Assessment of Hypernasality and Audible Nasal Emission in Children With Repaired Cleft Palate. Cleft Palate Craniofac J. 2015;52(6):660-70.

42. Pereira VJ, Tuomainen J, Lee KYS, Tong MCF, Sell DA. A perceptual outcome measure of velopharyngeal function based on the Cleft Audit Protocol for Speech-Augmented (CAPS-A VPC-Sum): Validation through a speech osteotomy study. Int J Lang Commun Disord. 2021;56(4):754-67.

43. Smarius BJA, Haverkamp S, de Wilde H, van Wijck-Warnaar A, Mink van der Molen AB, Breugem CC. Incidence of cleft-related speech problems in children with an isolated cleft lip. Clin Oral Investig. 2020.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

44. Vallino LD, Zuker R, Napoli JA. A study of speech, language, hearing, and dentition in children with cleft lip only. Cleft Palate Craniofac J. 2008;45(5):485-94.

45. Teele DW, Klein JO, Chase C, Menyuk P, Rosner BA. Otitis media in infancy and intellectual ability, school achievement, speech, and language at age 7 years. Greater Boston Otitis Media Study Group. J Infect Dis. 1990;162(3):685-94.

46. Ruegg TA, Cooper ME, Leslie EJ, Ford MD, Wehby GL, Deleyiannis FWB, et al. Ear Infection in Isolated Cleft Lip: Etiological Implications. The Cleft palate-craniofacial journal : official publication of the American Cleft Palate-Craniofacial Association. 2017;54(2):189-92.

to beet terien only

Protected by copyright, including for uses related to text

data mining, Al training, and similar technologies

FIGURE LEGENDS

Figure 1. Speech outcomes of ICHOM Cleft Lip and Palate and field test data

PROMs filled out by the patients are colored in purple, and grey if filled out by the parent(s). Clinical outcome measures are colored in yellow.

*Data derived from the field test only in these cohorts, as there is no measurement moment included in ICHOM Cleft Lip and palate in these age-groups.

Figure 2. Correlations in patients with CP and CL(A)P

All correlations in both cleft types appeared significant (P<0.05), except for the correlation between the PCC and CLEFT-Q SDistress in patients with CL(A)P (p = 0.285). Note that VPC is inversely scored (higher numbers correspond to worse outcomes), thus accounting for the negative correlations with the other scales.

Figure 3: cross-sectional trend analyses of the age-groups

Analyses are presented per outcome measure, per cleft type.

Figure 4 Overview of the new proposed ICHOM Cleft Lip and Palate set concerning speech assessment

Newly made recommendations are coloured in pink

*Suggestion for centres that have adequate resources to implement and are interested in research with speech oucomes.

Page 29	9 of 34			BMJ Open ft, inc		
1	5-7 years	8-9 years	10-13 years	14-16 years $\frac{1}{2}$	17-19 years	20-22 years
2				for us		
3 4 5	ICS	SFunction *	ICS	SFunction * regine 8	SFunction *	PCC
6 7 8	PCC	SDistress *	PCC	SDistress * to the second	SDistress *	VPC
9 10 11 12	VPC		VPC	nloaded fr perieur (A t and data		SFunction
13 14 15 16			SFunction	om http://t BES) . mining, Al		SDistress
17 18 19			SDistress	training,		
20 21 22				mj.com/ c and simili		
23 24 25 26				ar techno		
27 28 29				4, 2025 at logies.		
30 31 32				Agence		
33 34 35				Bibliogr		
36 37 38			For peer review only - http://b	niopen hmi com/site/about/quidelites	vhtml	
39 40			For peer review only - nttp.//bl		AH(H)	

BMJ Open





270x119mm (300 x 300 DPI)





BMJ Open

3	
1	
4	
5	
6	
7	
8	
9	
10	
10	
11	
12	
13	
14	
15	
16	
17	
17	
18	
19	
20	
21	
22	
22	
23	
24	
25	
26	
27	
28	
20	
29	
30	
31	
32	
33	
34	
25	
22	
36	
37	
38	
39	
40	
⊿1	
40 40	
42	
43	
44	
45	
46	
Δ7	
77 70	
4ð	
49	
50	
51	
52	
53	
55	
54 57	
55	
56	
57	
58	
59	
~ ~	

60

	Income	Beta	SE	P-value	Difference between groups	P-value
ed	High (ref)	72.10	0.51		High VS Middle	<0.001*
elat s	Middle	-16.39	1.58	<0.001*		
res Les						
ing	High (ref)	72.10	0.51		High VS Low	<0.001*
D	Low	-13.08	1.68	<0.001*		
Spe						
•,					Middle VS Low	0.285
	High (ref)	70.87	0.52		High VS Middle	<0.001*
sch ction	Middle	-12.32	1.64	<0.001*		
pee	High (ref)	70.87	0.52		High VS Low	<0.001*
Ч		12 10	1 71	<0.001*		<0.001
	LOW	-13.18	1./1	<0.001		
					Middle VS Low	0.923

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

Univariate analyses. Outcome scores of patients deriving from countries with a high income status were taken as reference. *Significant difference

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

Appendix 2	: descriptives	for subgroup	analyses
------------	----------------	--------------	----------

	CP (n=762)	CL(A)P (n=1379)	Total (n=2141)
Gender			
Male	308 (40.4%)	904 (65.6%)	1212 (56.6%)
Female	454 (59.6%)	475 (34.4%)	929 (43.4%)
SDistress			
Number	517 (67.8%	1068 (77.4%)	1585 (74.0%)
Missing	245 (32.2%)	311 (22.6%)	556 (26.0%)
SFunction			
Number	519 (68.1%)	1076 (78.0%)	1595 (74.5%)
Missing	243 (31.9%)	303 (22.0%)	546 (25.5%)
VPC			
Number	173 (22.7%)	213 (15.4%)	386 (18.0%)
Missing	589 (77.3%)	1166 (84.6%)	1755 (82.0%)
ICS			
Number	169 (22.2%)	228 (16.5%)	397 (18.5%)
Missing	593 (77.8%)	1151 (83.5%)	1744 (81.5%)
PCC score			
Number	152 (19.9%)	192 (13.92%)	344 (16.1%)
Missing	610 (80.1%)	1187 (86.1%)	1797 (83.9%)

Descriptives of patients deriving from high income countries, including all patients from the ICHOM centers.
Appendix 3: maximum scores and VPC overview

		Number of measurments with max score	Total measurements	
СР	SDistress	18.8% (97)	517	
	SFunction	19.8% (103)	519	
	ICS	29.0% (49)	169	
CL(A)P	SDistress	15.2% (162)	1068	
	SFunction	15.9% (172)	1076	
	ICS	18.0% (41)	228	
Maximum scores				

		СР			CL(A)P	
	Competent	Marginally	Incompetent	Competent	Marginally	Incompetent
	(0)	incompetent	(2)	(0)	incompetent	(2)
		(1)			(1)	
5 years	43 (50%)	21 (24.4%) 🧹	22 (25.6%)	11 (15.5%)	17 (23.9%)	43 (60.6%)
12 years	39 (65%)	19 (31.7%)	2 (3.3%)	46 (52.3%)	28 (31.8%)	14 (15.9%)
22 years	11 (40.7%)	13 (48.1%)	3 (11.1%)	26 (48.1%)	19 (35.2%)	9 (16.7%)

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

BMJ Open

What is the optimal assessment of speech? A multicenter, international evaluation of speech assessment in 2500 patients with a cleft.

Journal:	BMJ Open
Manuscript ID	bmjopen-2023-071571.R1
Article Type:	Original research
Date Submitted by the Author:	04-Oct-2023
Complete List of Authors:	Ombashi, Saranda; Erasmus Medical Center, plastic and reconstructive surgery Kurniawan, Melissa; Erasmus Medical Center, plastic and reconstructive surgery Allori, Alexander; Duke University Hospital & children's health center, Plastic, Maxillofacial & Oral Surgery Sharif-Askary, Banafsheh; Duke University Hospital & children's health center, Plastic, Maxillofacial & Oral Surgery Rogers-Vizena, Carolyn; Boston Children's Hospital, Plastic and Oral Surgery Koudstaal, Maarten; Erasmus MC Sophia Children Hospital, Department of Oral and Maxillofacial Surgery Franken, Marie-Christine; Erasmus Medical Center, Ear, Nose and Throat, Speech Therapy Mink van der Molen, Aebele; University Medical Center Utrecht, Department of Plastic Surgery Mathijssen, Irene; Erasmus MC Sophia Children Hospital, Department of Plastic and Reconstructive Surgery Klassen, Anne; McMaster University, Pediatrics Versnel, Sarah; Erasmus MC Sophia Children Hospital, plastic and reconstructive surgery
Primary Subject Heading :	Paediatrics
Secondary Subject Heading:	Dentistry and oral medicine, Evidence based practice, Paediatrics, Patient-centred medicine, Surgery
Keywords:	ORAL & MAXILLOFACIAL SURGERY, Paediatric head & neck surgery < PAEDIATRIC SURGERY, Paediatric oral & maxillofacial surgery < PAEDIATRIC SURGERY, Paediatric plastic & reconstructive surgery < PAEDIATRIC SURGERY, PLASTIC & RECONSTRUCTIVE SURGERY, Quality of Life
	·

1	
2	
3	
4	Manuscripts
6	Manuscripts
7	
8	
9	
11	
12	
13	
14	
15	
17	
18	
19	
20 21	
22	
23	
24	
25 26	
27	
28	
29	
30 31	
32	
33	
34	
35 36	
37	
38	
39	
40 41	
42	
43	
44	
45 46	
47	
48	
49	
50 51	
52	
53	
54	
55 56	
57	
58	
59	For peer review only - http://bmionen.bmi.com/site/about/quidelines.ybtml
00	

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our <u>licence</u>.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which <u>Creative Commons</u> licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

What is the optimal assessment of speech? A multicenter, international evaluation of speech assessment in 2500 patients with a cleft.	
S. Ombashi ^{1,2} , M. Kurniawan ^{1,} , A.C. Allori ³ , B. Sharif-Askary ³ , C.R. Rogers-Vizena ⁴ , M.J. Koudstaal ^{2, 5, 6} , M.C. Franken ⁷ , A.B. Mink van der Molen ^{2,8} , I.M.J. Mathijssen ^{1,2,6} , A.F. Klassen ⁹ , S.L. Versnel ^{1,2,6}	
¹ Department of Plastic and Reconstructive Surgery, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands	
² European Reference Network for rare and/or complex craniofacial anomalies and ear, nose and throat disorders	
³ Department of Plastic, Maxillofacial & Oral Surgery, Duke University Hospital & Children's Health Center, Durham, North Carolina, United States of America	
⁴ Department of Plastic and Oral Surgery, Boston Children's Hospital, Boston, Massachusetts, United States of	
5 Department of Oral and Maxillofacial Surgery, Special Deptal Care and Orthodoptics, Erasmus University Medical	
Center Botterdam The Netherlands	
⁶ Dutch Craniofacial Center, Frasmus University Medical Center, Rotterdam, The Netherlands	
⁷ Department of Ear, Nose and Throat, Speech Therapy, Frasmus University Medical Center, Rotterdam, The	
Netherlands	
⁸ Department of Plastic and Reconstructive Surgery, University Medical Center Utrecht. Utrecht. The Netherlands	
⁹ Department of Pediatrics, McMaster University, Hamilton, Canada	
Numer of words: 4436	
Number of tables: 2	
Number of figures: 3	
Number of appendices: 4	
Corresponding Author: S. Ombashi, MD. Dr. Molewaterplein 40, 3015 GD Rotterdam, Erasmus University Medical	
Center; Tel: +316-31989928; E-mail: <u>s.ombashi@erasmusmc.nl</u>	
ABSTRACT	
Objectives	
Speech problems in patients with a cleft palate are often complex and multifactorial. Finding the optimal	
way of monitoring these problems is challenging. ICHOM(International Consortium of Health Outcomes	
1	
For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	

2			
3 4	33	Measurement) developed a set of standardized outcome measures at specific ages for patients with a	
5 6	34	cleft lip and/or palate, including measures of speech assessment. This study evaluates the type and	
7 8	35	timing of speech outcome measures currently included in this ICHOM. Additionally, speech assessments	
9 10 11	36	in other cleft protocols and initiatives are discussed.	
12 13	37		
14 15	38	Design, setting and participants	
16 17	39	An international, multicenter study was set up including centers from the United States and The	
18 19 20	40	Netherlands. Outcomes of clinical measures and Patient Reported Outcome Measures(PROMs) were	
20 21 22	41	collected retrospectively according to the ICHOM set. PROM data from a field test of the CLEFT-Q were	
23 24	42	collected, including participants from countries with all sorts of income statuses, to examine the value of	
25 26	43	additional moments of measurement that are used in other cleft initiatives.	
27 28 29	44	Data from 2500 patients was included. Measured outcomes contained univariate regression analyses,	
30 31	45	trend analyses, T-tests, correlations and floor and ceiling effects.	
32 33	46		
34 35	47	Results	
36 37 38	48	PROMs correlated low to moderate with clinical outcome measures. Clinical outcome measures	
39 40	49	correlated low to moderate with each other too. In contrast, two CLEFT-Q scales correlated strongly	
41 42	50	with each other. All PROMs and the Percent Consonants Correct(PCC) showed an effect of age. In	
43 44	51	patients with an isolated cleft palate, a ceiling effect was found in the intelligibility in context scale(ICS).	
45 46 47	52		
48 49	53	Conclusion	
50 51	54	Recommendations for an optimal speech outcome assessment in cleft patients are made.	
52 53	55	Measurement moments of different cleft protocols and initiatives are considered in this	
54 55 56	56	proposition. Concerning the type of measures, adjustment of the current PCC score outcome	
57 58		2	
59 60		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	

57	seems appropriate. For centres with adequate resources and specific interest in research,
58	translation and validation of an upcoming tool, the CAPS-A, is recommended.
59	
60	Strengths and limitations of this study
61	International, multicenter setting
62	Data analyses per cleft type and age group
63	Cross-sectional data analyses
64	• 2500 participants with a cleft
65	Evaluating both PROMs and clinical outcome measures
66	
67	
07	
	For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml
	57 58 59 60 61 63 64 65 66 67

3	
4	
5	
6	
7	
, 0	
0	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
22 22	
2J 2∕I	
∠4 2⊑	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
30	
10	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
50	
5/	
58	
59	
60	

68	INTRODUCTION
69	A cleft lip and/or palate (CL/P) is the most common congenital craniofacial anomaly, with varying
70	incidence rates among Asians (1:500), Caucasians (1:1000) and patients of African descent (1:2500) (1-
71	4). Causes of a cleft are multifactorial, as both environmental and genetic factors have been reported(4).
72	Clefts can be categorized in multiple classification systems, of which a commonly used classification
73	system includes 4 cleft types: a cleft lip (CL); a cleft lip and alveolus (CLA); a cleft palate (CP); a cleft lip,
74	alveolus and palate (CL(A)P)(5). In addition, clefts can occur unilaterally or bilaterally(3).
75	Due to the facial defects, functional and appearance-related problems can occur, of which the extent
76	may depend on the cleft type; the severity of the cleft; and the coping of the individual and his/her
77	environment(6). Functional problems such as speech problems, hearing impairment and orodental
78	problems are often reported. As a result of the latter, difficulties with eating, drinking and breathing can
79	occur as well (5, 7).
80	Given the broad range of problems a patient with a cleft may have to face, treatment of patients with
81	CL/P is ideally done in a specialized and multidisciplinary cleft team in which speech therapists;
82	maxillofacial and plastic surgeons; otolaryngologists; pediatricians; psychologists; orthodontists;
83	geneticists; and specialized nurses are involved(7). Treatment and monitoring patients with CL/P
84	consists of multiple surgical interventions to close the defect and to improve appearance if the patient
85	desires so. Follow up of hearing function is indicated in case of a cleft palate, and placement of moppets
86	is regularly done if necessary. Furthermore, psychological guidance is often indicated while the child
87	grows up. Moreover, speech monitoring and long-term, intensive speech therapy is often necessary to
88	improve the eligibility of the child(5, 7).
89	
90	The development of speech is often complex in patients with a cleft palate (with or without a cleft lip,
91	CP±L). Persistent velopharyngeal incompetence, residual fistula, adenoid atrophy, surgical intervention

1 ว			
2 3 4	92	and hearing problems influence speech disorder severity in this population(8-12). Speech problems in	
5 6	93	patients with CP±L can have a large impact on an individual's life, as proper speech skills play an	
7 8 0	94	essential role in activities, social functioning and participation in society(13). Many treatment pathways	5
9 10 11	95	are focused on speech improvement to ameliorate Quality of Life (QoL)(14). Logically, speech	
12 13	96	assessment is an important parameter in cleft care.	
14 15	97		
16 17 19	98	However, no consensus has been reached regarding best diagnostic speech outcome measures and the	eir
19 20	99	timing in this population(5). Developing scientifically solid instruments to assess speech in an objective	
21 22	100	manner is complicated, because listener's perception of speech deficits, even by experts, may differ	
23 24	101	substantially(15). An additional challenge is systematic assessment of the patient's perspective, which i	S
25 26 27	102	essential to include due to the impact of speech problems on the individual(16). Although widely-	
27 28 29	103	accepted agreement seems essential for improvement of cleft care, finding consensus is complex,	
30 31	104	especially since speech outcomes should be comparable between different languages to facilitate	
32 33	105	international collaboration.	
34 35	106		
30 37 38	107	Recently, the International Consortium for Health Outcomes Measurement (ICHOM) developed the	
39 40	108	ICHOM Standard Set for Cleft Lip and Palate (ICHOM Standard Set), with different pathways for varying	
41 42	109	cleft types(5). Based on patient and expert consensus, a minimal, accessible set of outcome measures	
43 44	110	was established to enable benchmarking between cleft centers in a systematic manner. For speech	
45 46 47	111	assessment, an outcome set was included with both clinical measures and Patient Reported Outcome	
47 48 49	112	Measure (PROMs), being the patient's and parent's perspectives.	
50 51	113		
52 53	114	So far, the selected standardized speech outcome measures and their timing have not been evaluated.	
54 55 56	115	As an increasing number of centers are implementing this set, it is important to critically evaluate and	
50 57 58			5
59 60		For peer review only - http://bmiopen.bmi.com/site/about/quidelines.xhtml	2
00			

1		
2 3 4	116	optimize this ICHOM Standard Set. Three centers, the Boston Children's Hospital (Boston, USA), Duke
5 6	117	University Hospital (Durham, USA), and the Erasmus Medical Center (Rotterdam, The Netherlands),
7 8	118	started clinical implementation and an international collaboration in 2015. The overarching aim of this
9 10 11	119	collaboration is to share data and knowledge obtained by using the set in standard care. Additionally,
12 13	120	they collaborate with McMaster University (Hamilton, Canada), who developed and field tested the
14 15	121	CLEFT-Q questionnaire, of which many scales are included in the ICHOM Standard set.
16 17	122	
18 19	123	The objective of this study was to evaluate the current standardized speech outcome measures of the
20 21 22	124	ICHOM Standard Set for patients with CP±L. More specifically, the value of every speech outcome
23 24	125	measure was examined, as well as the best age intervals for assessment of these outcome measures. In
25 26	126	addition, other speech assessment tools are discussed. Finally, recommendations are made for an
27 28	127	optimal and complete assessment of speech in patients with CP±L, that is efficient and accessible for all
29 30 31	128	cleft centers.
32 33	129	
34 35	130	
36 37		
38 39 40		
41 42		
43 44		
45 46		
47 48		
49 50		
51 52		
53 54 55		
55 56 57		
58 59		6
60		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

2 3 4	131	METHODS
5 6 7 8	132	Patient population
9 10 11	133	Three centers (Boston Children's Hospital, Duke University Hospital, and Erasmus Medical Center,) each
12 13	134	implemented the ICHOM Standard Set in 2015. All patients treated at these centers for a cleft palate
14 15	135	with a cleft lip/cleft alveolus (CL(A)P), or an isolated cleft palate (CP) who were assessed according to
16 17	136	the ICHOM Standard Set (age range 5-22 years), were included. In addition, another patient group
18 19	137	derived from an international field test of the CLEFT-Q, by McMaster University(17). According to the
20 21 22	138	age cut-off of the ICHOM Standard Set, only outcomes from CLEFT-Q scales of field test patients with a
22 23 24	139	CP±L up to 22 years old were included in the current study(appendix 1). Patients from the participating
25 26	140	centers were excluded in case they could not sufficiently speak or write the language native to the
27 28 29 30 31 32 33 34 35 36 37	141	center's country .
	142	
	143	Patient and public involvement
	144	In the development of the ICHOM Standard Set, patients were actively involved. The ICHOM Standard
38 39	145	Set was implemented in each center as part of regular clinical care. Data was pseudonymized and
40 41	146	collected retrospectively, and ethical approval was obtained to do so without explicit consent from each
42 43 44	147	patient and parent. Results of this study may be of use to further improve the currently used ICHOM
45 46	148	Standard Set, and therewith regular clinical care.
47 48	149	
49 50	150	
51 52 53	151	Outcome measures
54 55	152	PROMs
56 57		-
58 59 60		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

3 4	1
5 6	1
7 8	1
9 10	1
11 12	1
13 14	1
15 16	1
17 18	1
19 20 21	L
21 22 22	1
23 24 25	1
25 26 27	1
27 28 20	1
30 31	1
32 33	1
34 35	1
36 37	1
38 39	1
40 41	1
42 43	_
44 45	_
46 47	1
48 49	1
50 51	1
52 53	
54 55	
56 57	
58	
59 60	

1 2

153	<u>CLEFT-Q Scales</u>
154	The CLEFT-Q is developed specifically to assess QoL from the patient's perspective in patients with a
155	CP±L. A literature review, patient interviews and psychometric testing, established the final content of
156	the scales, which covers several overarching domains(18-20). Speech is assessed through two scales,
157	each covering a different domain. Both scales have 3 response options for each item (always;
158	sometimes; never); a lower score equals a worse outcome. Completing the scales can be done online; it
159	will take the patient several minutes.
160	Speaking-Related Distress (SDistress) is part of the psychosocial domain. The scale contains 10 items that
161	relate to the psychosocial part of speech difficulties, like nervousness or frustration(20).
162	Speech Function (SFunction) focuses on the functional speech difficulties that patients themselves
163	identify, for example the ability to say certain letters or words. The scale consists of 12 items that belong
164	to the facial function domain(20).
165	
166	Intelligibility in Context Scale(ICS) is a measure that assess the intelligibility of the child. It is a 7-item,
167	parent-reported questionnaire designed to be scored by speech pathologists. The score indicates a
168	child's level of functional intelligibility, by assessing the degree to which the speech of the patient is
169	understood by different communication partners. The total score is calculated by the averages of the
170	items completed. ICS appeared to be a valid and reliable tool for children with speech disorders, (21,
171	22), but not specifically designed nor validated for patients with CP±L(23). It is available in several
172	languages, and normative data exists for English-speakers(24, 25).
173	
174	Clinical outcome measures

1			
2 3 4	175	Percent Consonants Correct(PCC) is developed to detect speech sound errors. PCC scores are calculated	
5 6	176	by using a standard, crossecitonally translated set of words that include all speech problems children	
/ 8 9	177	with CP±L often tend to have.	
10 11	178	In case of any problems, their severity can be categorized: PCC scores of 85-100% indicate mild to no	
12 13	179	problems; scores of 65-84.9% indicate mild-moderate problems; scores of 50-64.9% indicate moderate-	
14 15	180	severe problems; and scores <50% indicate severe problems(20). PCC is suitable for usage in patients	
16 17 18	181	with CP±L when assessed by well-trained clinicians(8).	
19 20	182		
21 22	183	<u>Velopharyngeal Competence rating(VPC)</u> discriminates between three categories: 'competent',	
23 24	184	'marginally incompetent' and 'incompetent'. The outcome is determined by the speech therapist based	
25 26 27	185	on the PCC test and spontaneous speech. In case of any clinical evidence of minor problems regarding	
27 28 29	186	the competence, VPC was categorized as 'marginally incompetent'. When clinically significant problems	
30 31	187	were detected, suggesting surgical management and/or speech therapy, VPC was categorized as	
32 33	188	'incompetent'. Prior studies found VPC to be suitable as a first clinical choice for the assessment of	
34 35 36	189	velopharyngeal dysfunction and is recommended for both clinical follow-up and research(26).	
37 38	190		
39 40	191	Data collection	
41 42	192	All participating centers obtained ethical approval for the current study from their local ethics	
43 44 45	193	committees.	
46 47	194	Data was collected restrospectively over a 6 year period (2015-2020) and extracted from the electronic	
48 49	195	patient files in 2018 and 2020 (as a data update). Both video and audio records were used for the	
50 51	196	evaluation of the clincal outcome measures. During the data collection period, the included center	
52 53 54 55 56	197	cooperated together, and regular meetings were held (both online and live).	
57 58		9	
59 60		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	

1

60

2		
3 4	198	According to the ICHOM protocol, both CLEFT-Q Speech scales were assessed at ages 12 and 22
5 6	199	years(appendix 1). Both PCC and VPC were scored at ages 5, 12 and 22 years, and ICS at ages 5 and 12
7 8	200	years.
9 10 11	201	The field test cross-sectionally collected data from patients with a cleft across 12 different countries
12 13	202	with different income-statuses(17). As 8 years is the minimum age to complete the CLEFT-Q, both
14 15	203	CLEFT-Q Speech scales from field test patients with CP±L from 8-22 years old were included (appendix
16 17	204	1).
18 19 20	205	Income status of the country according to the World Bank Classification was made within the field test.
20 21 22	206	Data from the ICHOM centers were all categorized as deriving from high-income countries.
23 24	207	Baseline characteristics that were collected included gender, type of cleft, and age at the time of
25 26	208	assessment.
27 28	209	
29 30 31	210	Data analysis
32 33	211	Data was analysed in R-studio, a free software environment for statistical computing and graphics(27).
34 35	212	Psychometric validation of the SDistress and SFunction confirmed suitability to use a 0 to 100 scale
36 37	213	deriving from the sum scores for analysis(17).
38 39 40	214	For analysis of the ICS questionnaire, the average score of the seven items was used. VPC was used as
40 41 42	215	an ordinal variable, whereas PCC scores were expressed as proportions.
43 44	216	All participating ICHOM centers are high-income countries, whereas part of the field test data was
45 46	217	collected in upper middle and lower middle income countries. To prevent possible influence of income-
47 48 40	218	status on the outcomes, univariate regression analyses were used to examine differences in outcome
49 50 51	219	scores of the SDistress and SFunction before further analyses. Data was categorized according to the
52 53	220	income status of the country where the data had been collected.
54 55		
56 57		
58 59		10

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1		
2 3 4	221	In order to examine the added value of each PROM and clinical outcome measure in regard with the
5 6	222	other measures, correlations were examined between the PROMs; between the clinical outcome
7 8	223	measures; and between the PROMs and the clinical outcome measures. Pearson correlations were used,
9 10 11	224	and outcomes were analysed per cleft type. Correlations were considered strong in case r>0.7;
12 13	225	moderate between r=0.5 and R= 0.7; and weak in case r<0.5.
14 15	226	Analyses within and between different age-groups were done to explore whether the current outcome
16 17	227	measures are assessed at the optimal age-points, and whether additional measurement moments are
18 19 20	228	indicated either for PROMs or clinical outcome measures. Therefore, not only time-points of the ICHOM
20 21 22	229	protocol were included in analyses. As CLEFT-Q outcome scores of all ages between 8-22 years were
23 24	230	included from the field test data, time-points used by other large initiatives as Eurocleft, Scandcleft and
25 26	231	Americleft were considered as well (28-30). Doing so, the following age-groups were set up: 5-7 years; 8-
27 28 20	232	9 years; 10-13 years; 14-16 years; 17-19 years, 20-22 years(app. 2).
29 30 31	233	Per age-group, possible differences between scale scores were examined with independent T-tests.
32 33	234	Bonferroni correction was applied for multiple testing. Trend analyses were performed to identify
34 35	235	potential problems in specific age-groups.
36 37 28	236	Floor- and ceiling effects were examined to identify the suitability of the outcome measures in our
39 40	237	population. A floor or ceiling effect is seen when a considerable amount of the outcome scores are
41 42	238	either scored the best (in this case a maximum score, thus a ceiling effect), or the worse (in this case a
43 44	239	minimum score, thus a floor effect). Both effects result in a truncated distribution of the outcomes on
45 46 47	240	either side of the scale (31, 32). Minimum and maximum score outcomes of all PROMs and the PCC were
47 48 49	241	evaluated. A percentage of 20% or more of the patients scoring the minimum or maximum outcome
50 51	242	score was considered as a ceiling effect. In VPC, the outcome distributions were examined.
52 53	243	
54 55	244	
56 57 58		11
59		
60		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Page	14 of	55
	Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.	MJ Open: first published as 10.1136/bmjopen-2023-071571 on 28 December 2023. Downloaded from http://bmjopen.bmj.com/ on June 14, 2025 at Agence Bibliographique de I

12

2						
3 4	245	RESULTS				
5 6	246					
7 8	247 Characteristics of the included participants					
9 10 11	248	A total of 2500 patients w	ere included in the s	tudy; 1723 derived	from the field test,	and 777 from the
12 13	249	ICHOM centers (table 1).				
14 15	250					
16 17	251 252	Table 1: Demographics an	d phenotypes			
18 19	252	Descriptives total				
20 21	200		Field test (n = 1723)	ICHOM (n = 777)	Overall (n = 2500)	
22 23		Sex				
24		Male	981 (56,9%)	444 (57,1%)	1425 (57.0%)	
25 26		Female	742 (43 1)	333 (42 9%)	1075 (43.0%)	
27			, 12 (13.1)	333 (42.570)	1073 (43.070)	
28 29			517 (30.0%)	201 (28 7%)	818 (32 7%)	
30			1206 (70.0%)	476 (61.2%)	1692 (67.2%)	
31 32			1208 (70.0%)	470 (01.3%)	1082 (07.5%)	
33			12(4/70.20/)	777 (100%)	21.41 (85.6%)	
34 35		Hign Income	1364 (79.2%)	777 (100%)	2141 (85.6%)	
36		Up Middle income	199 (11.5%)	0 (0%)	199 (8.0%)	
37 38	254	Low Middle Income	160 (9.3%)	0 (0%)	160 (6.4%)	
39	254					
40 41	255					
42						
43	256	There were slightly more males than females, and relatively more patients with a CL(A)P than with a CP.				
44 45 46	257	Significant differences between countries with a High and Upper-/Lower-Middle income status of the				
47 48	258	field test were found and results are shown in appendix 2.				
49 50	259	Therefore, further analyses were done only with the patient population of the field test deriving from				
51 52 53	260	countries with a high income status, like all participating ICHOM Cleft centers (n = 2141). The subgroup				
55 54 55	261	characteristics are include	d in appendix 3.			
56 57						
58						12
59 60		For peer r	eview only - http://bn	njopen.bmj.com/site	/about/guidelines.xh	tml

1								
2 3 4	262	Associatio	ons between the	outcome measur	es			
5 6	263	Correlations between all outcome measures (clinical and PROMs), in both cleft types (CP and CL(A)P),						
7 8 0	264	appeared	significant (P<0.0)5), except for the	correlation betwee	en PCC and SDistr	ess (p = 0.285)(fig. 1)	
9 10 11	265	<u>Correlatio</u>	<u>n PROMs:</u> The SI	Distress and SFund	tion showed a stro	ng correlation in _l	patients with CP (r =	
12 13	266	0.76) and	a moderate corre	elation in patients	with CL(A)P (r = 0.6	58). The ICS and S	Function correlated	
14 15	267	strongly (r	[.] = 0.73) in patier	its with CP, and m	oderately (r = 0.64)	in patients with	CL(A)P; whereas the I	CS
16 17 18	268	and the SI	Distress correlate	d moderately in p	oatients with CP (r =	0.52) and weak i	n patients with CL(A)	P (r
19 20	269	= 0.47).						
21 22	270	<u>Correlatio</u>	n clinical outcom	es: VPC and PCC v	were (negatively) m	oderately correla	ited in both cleft type	!S
23 24 25	271	(r = -0.62 a	and r = -0.67 in p	atients with CP ar	d CL(A)P respective	ely).		
25 26 27	272	Correlation PROMs and clinical outcomes: Moderate correlations were found between the PCC and ICS						
28 29	273	in patients with CP (r = 0.64), and in patients with CL(A)P (r= 0.5). VPC and ICS had a (negative) weak						
30 31	274	correlation in patients with CP (r = -0.49), and CL(A)P (r = -0.43). The SDistress and SFunction were						
32 33	275	weakly correlated with VPC and PCC (negatively) in both cleft types (fig. 1).						
34 35 36	276	Comparing outcome measures between age-groups						
37 38	277	SDistress and SFunction: showed the highest mean outcome (i.e. the most favourable ratings) in the						
39 40	278	age-group of 14-16 years old. From thereon, a slightly downward trend is seen (fig. 2). In both CLEFT-Q						
41 42	279	Speech scales the lowest mean outcome scores were found in the age-group of 8-9 years old, which was						
43 44 45	280	significantly different in comparison with the other age-groups in patients with CL(A)P (p<0.05, table 2).						
45 46 47	281							
48	282	Table 2 N	/lean outcomes	per age-group,	per cleft type			
49			Age	SDistress	SFunction	ICS	PCC	
50 51			categories					
52		. u	5 – 7			4 08 (0 73)	72 49 (31 14)	
53		left lat	<u> </u>	70 11 (20 61)	66 12 (22 12)	1.00 (0.75)	, 2.75 (31.17)	
54		Pa	0 - 9	70.44 (20.04)				

92.02 (11.48)1

72.56 (21.52)

4.45 (0.62)¹

74.84 (19.91)

10 - 13

59

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

	14 – 16	75.10 (22.77)	74.66 (24.32)		
	17 – 19	70.59 (16.92)	73.00 (19.02)		
	20 – 22	70.27 (21.93)	71.69 (24.98)		90.92 (15.33) ¹
	5 – 7			4.03 (0.50)	46.06 (25.59)
	8-9	65.32 (21.61)	64.75 (20.16)		
ie in	10 - 13	73.16 (20.62) ¹	69.57 (20.96)	4.32 (0.55) ¹	76.25 (22.63) ¹
eft I alat	14 – 16	72.81 (19.66) ¹	74.22 (19.14) ¹		
E E	17 – 19	71.82 (19.67) ¹	71.41 (20.79) ¹		
	20 – 22	71.49 (18.00)	72.26 (18.40) ¹		86.83 (16.13) ^{1,2}

The variables in the table present the mean outcomes of the SDistress, SFunction, ICS and PCC. Outcomes were categorized into cleft type and age group.

p < 0.05 compared to 5-7 or 8-9 age

P < 0.05 compared to 10 - 13 age

ICS : Both patient groups showed a significant difference between the two age-groups 5 and 12 years,

ICS was significantly lower at 5 years than at 12 years in both cleft types (tab. 2).

PCC: Observing the trends in the clinical outcome measures, an upward trend regarding PCC score was

seen (fig. 2). In both cleft types, PCC scores differed significantly between the age-groups(tab. 2).

VPC: In the age-group of 5 years, 25.6% of the patients with CP and 60.6% of the patients with CL(A)P

were scored as incompetent. In 22 year-olds, this percentage was 11.1% in patients with CP and 16.7%

in patients with CL(A)P.

No floor effects were found in any of the PROMs. In patients with CP, the ICS showed a ceiling effect

(29.0%, n = 169). No ceiling effects were observed in patients with CL(A)P. An overview of all maximum

scores and the VPC score distribution is shown in appendix 4.

1		
2 3 4	299	DISCUSSION
5 6	300	
7 8	301	Evaluation of
9 10 11	302	All correlation
11 12 13	303	both the SDist
14 15	304	and SDistress
16 17	305	different appe
18 19	306	context, lookii
20 21 22	307	is supported b
22 23 24	308	in the CL(A)P ք
25 26	309	their speech f
27 28	310	latter, distress
29 30 31	311	indication of '
32 33	312	yet (7 years ar
34 35	313	
36 37	314	The PROMs sh
38 39 40	315	correlation that
40 41 42	316	PROMs appea
43 44	317	clinical outcon
45 46	318	outcome mea
47 48 40	319	speaking situa
49 50 51	320	probably be in
52 53	321	aspects of SDi
54 55		
56 57		
58		
59 60		
47 48 49 50 51 52 53 54 55 56 57 58 59 60	319 320 321	speaking situa probably be ir aspects of SDi

15

n of the value of the current ICHOM speech outcome measures tions between PROMs were moderate, except for the strong correlation of the SFunction with Distress and the ICS in patients with a CP. The fact that the correlation between the SFunction ess is stronger in patients with CP than in patients with CL(A)P, suggests that the visibly appearance in patients with CL(A)P plays a significant role in SDistress as well; in a social poking differently may cause additional or more distress besides having speech problems. This ed by our finding that the ICS correlated moderately with SFunction, but weakly with SDistress A)P group. Parent reported speech intelligibility correlated higher to children's self report of ch function, than it did to the speech distress the children themselves experience. In the ress about appearance could be included. This finding suggests that the ICS can give an of 'patient-reported' SFunction in young children who cannot complete a PROM themselves rs and younger). Is showed weak correlations with the clinical reported outcomes; except for the moderate n that was seen between the ICS and the PCC in both patient groups. Based on these findings, ppear to be of added value, as they provide different information than captured with the tcome measures included in the Standard Set. They add a unique dimension to speech measurement – a subjective dimension related to the patient's experiences with everyday situations. While clinical measures objectively appraise the quality of speech, they will be insufficient to adequately capture the more nuanced social, emotional, and psychological SDistress and SFunction. With this additional self-report and parental information, clinicians

Page 18 of 55

BMJ Open: first published as 10.1136/bmjopen-2023-071571 on 28 December 2023. Downloaded from http://bmjopen.bmj.com/ on June 14, 2025 at Agence Bibliographique de Enseignement Superieur (ABES)

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies

322	can more comprehensively explore the patients' problems concerning speech in order to find out
323	whether additional treatment or guidance is indicated.
27/	

Evaluation of the impact of age of assessment on measurement outcomes

In both CLEFT-Q Speech scales, the age-group of 8-9 years enholds the worst scores. Speech improvement due to speech therapy or late closure of the hard palate (in certain protocols around the age of 9 years when alveolar bone grafting is performed), might explain the higher, better scores in the age-groups of 10-13 and 14-16 years. In age-groups 17 and up however, CLEFT-Q scores appeared to decline whereas PCC scores improved. This finding suggests that (almost) adult patients with CP±L develop feelings of insecurity concerning their speech, although their speech sound production remains good, or even improves. This is in line with speech therapists' experiences in the outpatient clinic, where patients were seen in person at the age of 22, but not at age of 17-19. Quite often, when discussing outcomes of the CLEFT-Q scales as well as the PCC with the patient, (s)he reacted surprised when told that no (cleft-related) problems were present in their speech. Taking the lower CLEFT-Q scores in 8-9-, 17-19- and 20-22-year-olds that were found in the field test into consideration, additional assessment of a PROM at the age-groups of 8-9 (youngest age at which this PROM can be assessed) and 17-19 years old should be considered for implementation in the ICHOM Standard Set. Therewith, monitoring patients more closely will be enabled, and any concerns of patients with CP±L regarding their speech can be discussed timely. The two CLEFT-Q speech scales showed to capture overlapping information as they strongly correlate in patients with CP. Questions deriving from the SDistress are not measurable in any other manner, whereas SFunction from the patient's perspective might be less of added value for a PROM questionnaire. Therefore, implementation of the CLEFT-Q SDistress scale in patients with both cleft types is recommended at the age-groups of 8-9 years and 17-19 years (fig. 3).

1		
2 3 4	346	
5 6	347	A ceiling effect in ICS outcomes of patients with CP, without clear differences between average scores in
/ 8 9	348	patients with CP and CL(A)P, suggests that the group with CP contains a diverse population and severity
10 11	349	of the speech problems vary widely. Furthermore, since ICS is not specifically developed for a population
12 13	350	with CP±L, it is debatable whether this tool captures the information necessary to point out all relevant
14 15 16	351	speech problems in the patient group.
16 17 18	352	However, exclusion of ICS could mean that a large part of the speech problems in the population with CP
19 20	353	would remain undetected. Assessment at 5 and 12 years in patients with both cleft types, which is the
21 22	354	current timing in the ICHOM Standard Set, appears therefore appropriate despite the ceiling effect.
23 24	355	
25 26 27	356	Although VPC scores were relatively favourable in patients with CP, no changes regarding the
27 28 29	357	implementation of the VPC scores are recommended as the outcomes showed to vary. VPC can serve as
30 31	358	a suitable screening tool and outcomes are easily gathered by the observation of a clinician. Hence,
32 33	359	patient-burden is low and the tool efficiently detects any velopharyngeal problems.
34 35 26	360	
30 37 38	361	PCC scores that were found indicated speech sound problems especially in the younger age-groups of
39 40	362	the patients with CL(A)P. 22-year-olds with both cleft types showed mild speech sound problems in
41 42	363	general. Therefore time points as currently implemented in the ICHOM Standard Set appear adequate.
43 44 45	364	In contrast, the suitability of PCC-assessment in a cleft set focusing on standardized outcome measures
43 46 47	365	is still debatable, as inter- and intra-center reliability have not been investigated thoroughly in all
48 49	366	participating centers so far(15). Future research should include an examination of scoring and
50 51	367	interpreting PCC scores in different centers and/or different countries.
52 53	368	
54 55 56	369	
57 58		17
59 60		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1 2			
3 4	370	Future considerations regarding alternative speech outcome measures	
5 6	371	In order to establish an optimal cleft set for speech assessment, other standardized outcome measures	
7 8	372	should be considered. Based on clinical experience with ICHOM Standard Set, possible suggestions for	
9 10 11	373	additional outcome measures are discussed here.	
12 13	374		
14 15	375	Regarding PROMs for speech assessment in patients with CP±L, the CLEFT-Q scales seem to be the most	
16 17	376	suitable PROMs available. Their comprehensive psychometric examination and cross-cultural character	
18 19	377	make them accessible for all cleft centers that seek an efficient minimal cleft set that comprises all	
20 21 22	378	important speech parameters(17-19). The standardized approach for translation and validation of the	
23 24	379	CLEFT-Q questionnaire enables accessibility of the PROM even for centers that still need to translate the	
25 26	380	CLEFT-Q into their native language(33, 34). Another cleft-specific PROM is the Cleft Hearing and Speech	
27 28	381	Questionnaire (CHASQ). Whereas the psychometric properties of the CLEFT-Q were examined	
29 30 31	382	throughout Rasch measurement theory, classical test theory was used for the CHASQ(35). A recent	
32 33	383	cross-sectional questionnaire study that compared the CLEFT-Q with the CHASQ, found that the majority	1
34 35	384	of the patients with CP \pm L preferred the CLEFT-Q(35). Therefore, implementation of the CHASQ speech	
36 37	385	does not seem to be of added value in the current cleft set.	
38 39 40	386		
41 42	387	Besides the used VPC measure, a more elaborate variant exists, namely the VPC-Summary (VPC-Sum).	
43 44	388	This includes assessment of hypernasality, passive VPI symptoms and the transcriptions of active non-	
45 46	389	oral consonant errors(36). VPC-sum can either be reported as a score between 0 and 6, or as a	
47 48 40	390	dichotomized outcome (velopharyngeal competence or incompetence)(36).	
50 51	391	Calculation of the VPC-Sum is based on single-words, whereas VPC-rate is based on observation of	
52 53	392	spontaneous speech (37). VPC-Sum would be an interesting measure due to its efficiency, although it	
54 55	393	may not be achievable to implement VPC-Sum in all centers in the near future as only five different	
56 57 58		10	2
59 60		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	,

1 ว		
2 3 4	394	languages are currently available(31). Other alternatives such as nasopharyngoscopy or MRI are
5 6	395	invasive, expensive and enlarge the patient burden(38), and therefore not easy accessibility for all
7 8 0	396	centers.
9 10 11	397	
12 13	398	The currently implemented PCC lacks any categorization of consonant errors. The Eurocleft speech
14 15	399	group created a research protocol with a phonetic framework, which was used in six centers and five
16 17 19	400	different languages(39). It also included consonant production, but assessed on sentence level instead
19 20	401	of single words. It categorized into 3 groups (correct; almost correct; incorrect). Further division into 21
21 22	402	error categories that were sampled in five groups was done in case of incorrect consonants (Nasal
23 24	403	airflow; glottal realisations; alveolar deviations; sibilant deviations; other)(39). Moreover, general
25 26 27	404	speech quality was assessed concerning hyper- and hyponasality, and voice quality(40). Expert rating of
27 28 29	405	these outcomes requires periodic training of sufficient interrater reliability. However, it might be too
30 31	406	detailed for implementation in an efficient, clinically oriented cleft set. Therefore, we suggest to further
32 33	407	categorize the PCC score, although not as detailed as in the Eurocleft studies. Based on clinical
34 35 26	408	experience with the ICHOM Standard Set, it is recommended that speech pathologists report whether
30 37 38	409	any cleft related, phonological, or phonetic problems are detected.
39 40	410	
41 42	411	Another clinical outcome measure, the Great Ormond Street Speech Assessment 1998 (GOS.SP.ASS'98),
43 44 45	412	provides a comprehensive view of all speech associated features for patients with CP±L(41, 42). Its
45 46 47	413	suitability for inter-center comparison would make it interesting for the ICHOM Standard Set(5);
48 49	414	however, it is too detailed for clinical audit(43). In succession the Cleft Audit Protocol for Speech
50 51	415	Augmented (CAPS-A) was developed for cleft-related problems, and could be an alternative for PCC(44).
52 53	416	Seen its rigorous psychometric assessment, it fits well into a set that seeks standardized outcome
54 55 56	417	measures. The Americleft Speech Project found that an acceptable inter- and intrarater reliability can be
57 58		19
59 60		For peer review only - http://bmjopen.bmj.com/site/about/quidelines.xhtml

2	
- 3 4	41
5 6	41
7 8	42
9 10	42
11 12 13	42
14 15	42
16 17	42
18 19	42
20 21	42
22 23 24	42
25 26	42
27 28	42
29 30	43
31 32 33	43
33 34 35	43
36 37	43
38 39	43
40 41	43
42 43 44	43
45 46	43
47 48	43
49 50	43
51 52 53	44
53 54 55	44
56 57	
58	
59 60	
00	

418	achieved(43, 45). Furthermore, it is suitable for assessment in 5-years-olds, which enables detection of
419	speech problems at an earlier age(46). However, the CAPS-A is limited in types of statistical analyses due
420	to the scaling type used (equal appearing interval)(47). A more practical challenge concerning
421	implementing the CAPS-A would be the required training of all involved speech therapists, and the
422	amount of time the assessment takes (15 min.)(44). Moreover, the CAPS-A is developed and applicable
423	for English-speaking countries, necessitating translation and validation in other languages(43). The
424	CAPS-A is not ideal for centers interested in a minimal and efficient cleft set. However, centers with
425	experience and resources are highly recommended to implement this tool in order to promote further
426	international standardization of elaborate speech assessment in patients with CP±L(fig. 3).
427	Implementation of the CAPS-A would also enable the use of the recently developed and validated CAPS-
428	A-VPC-Sum score to reliably measure velopharyngeal function(48). Our suggestion for centers that
429	consider the implementation of the CAPS-A, is to assess it at the ages of 5-7, 10-13 and 20-22 years in
430	order to enable long-term follow-up.
431	
432	Limitations of the study
433	Data was analyzed cross-sectional. Longitudinal analyses to explore development of speech and for
434	benchmarking will be possible in the future, since data collection continues. Moreover, because this
435	study included data from the CLEFT-Q field test, a higher number of outcome data from the CLEFT-Q
436	scales was available for analyses than from the other outcome measures included in the ICHOM
437	Standard Set.
438	
439	
440	CONCLUSION
441	
	00
	20 For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

BMJ Open

From the current study, it can be concluded that the current ICHOM Standard Set is informative and efficient. PROMs were shown to be of added value, and the CLEFT-Q appeared to be the most suitable PROM. Therefore, continuation of collecting the current outcome measures and time points is recommended. Furthermore, a minor extension is suggested: In addition to the current time-points of assessment, it is recommended to implement the CLEFT-Q SDistress scale at the age of 8-9 and 17-19 as well. Further adjustments of the set could comprise an additional categorization of the PCC score, based on the framework of Eurocleft and adjusted for clinical usage. ACKNOWLEDGMENTS We would like to thank all children and caregives 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 speech pathologists from all our participating centers, that contributed tremendously by collecting the data consequently and by advising the research team on important clinical considerations

1 2		
2 3 4	459	Contributor statement
5 6	460	Saranda Ombashi : Data collection in Dutch Center; correspondence to other centers; data analyses;
7 8 9	461	writing and adjusted the manuscript based on advice of the other authors.*
10 11	462	Melissa S.I.C. Kurniawan: Data analyses, co-writer on the methods and results section of the
12 13	463	manuscript.*
14 15 16	464	Alexander C. Allori: Data collection at Duke University Hospital, advisor of project plan and manuscript.*
17 18	465	B. Sharif-Askary: Data collection at Duke University Hospital, advisor of project plan and manuscript.*
19 20	466	Carolyn R. Rogers: Data collection at Harvard University Hospital, advisor of project plan and
21 22	467	manuscript.*
23 24 25	468	Maarten J. Koudstaal: Writer of the Ethical Board Approval, Supervision of data analyses. Advisor and
23 26 27	469	supervisor on project plan and manuscript.*
28 29	470	M.C. Franken: Data collection at Erasmus Medical Cewnter, advisor of project plan and manuscript,
30 31	471	with special focus on the clinical content.*
32 33	472	A.B. Mink van der Molen: advisor of project plan and manuscript.*
34 35 36	473	Irene M.J. Mathijssen: Overall supervision of project plan and manuscript.*
37 38	474	Anne F. Klassen: Data collection Field Test Mc Master University, advisor on project plan and
39 40	475	manuscript.*
41 42	476	Sarah L. Versnel: Project leader, correspondence with all other centers. Supervision on project
43 44 45	477	plan, data analyses and manuscript.*
46 47	478	
48 49	479	*All authors approved the final version to be published and agreed to be accountable for all
50 51 52	480	aspects of the work in ensuring that questions related to the accuracy or integrity of any part of
53 54	481	the work are appropriately investigated and resolved
55 56		
57 58		22
59		

1		
2 3 4	482	Competing interests:
5	483	The CLEFT-Q is owned by McMaster University and the Hospital for Sick Children. Prof. Anne Klassen is a
/ 8 9	484	co-developer of the CLEFT-Q and would receive a share of any licence revenues as royalties based on her
10 11	485	institutions' inventor sharing policy if the CLEFT-Q was used in a for profit study.
12 13	486	The other authors have no disclosures.
14 15 16	487	
17 18	488	Funding
19 20 21	489	This project received no funding. However, S. Ombashi works as a PhD-student for an European
21 22 23	490	Union-funded Network, the 'Eureopean Reference Network for Craniofacial anomalies and Ear,
24 25	491	Nose and Throat disorders'. The results and findings from this project are of help in further
26 27 28	492	European alignment concerning standardized outcome measures in cleft care.
20 29 30	493	
31 32	494	Data sharing statement
33 34 35	495	A data sharing agreement was set up and signed by all participating centers that shared data for
36 37	496	this project.
38 39 40	497	
40 41 42	498	Ethics approval statement
43 44	499	All participating centers obtained ethical approval for the current study from their local ethics
45	500	committees (ethics committees of the Erasmus University Medical Center; Utrecht University
46 47	501	Medical Center; Duke University Hospital & Children's Health Center; Boston Children's
48	502	Hospital; McMaster University) IRB approval nr MEC-2016-156.
49 50		
50 51		
52		
53 54		
55		
56		
57 58		23
59		
60		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

2		
- 3 4	503	References
5	504 505	1. Hlongwa P. Levin J. Rispel LC. Epidemiology and clinical profile of individuals with cleft
6 7	506	lip and palate utilising specialised academic treatment centres in South Africa. PLOS ONE.
8	507	2019:14(5):e0215931.
9	508	2. Tanaka SA, Mahabir RC, Jupiter DC, Menezes JM, Updating the epidemiology of cleft lip
10	509	with or without cleft palate. Plast Reconstr Surg. 2012:129(3):511e-8e.
11 12	510	3. Tolarová MM, Cervenka J. Classification and birth prevalence of orofacial clefts. Am J
12	511	Med Genet. 1998;75(2):126-37.
14	512	4. Wang W, Guan P, Xu W, Zhou B. Risk factors for oral clefts: a population-based case-
15	513	control study in Shenyang, China. Paediatr Perinat Epidemiol. 2009;23(4):310-20.
16	514	5. Allori AC, Kelley T, Meara JG, Albert A, Bonanthaya K, Chapman K, et al. A Standard Set
17	515	of Outcome Measures for the Comprehensive Appraisal of Cleft Care. Cleft Palate Craniofac J.
19	516	2017;54(5):540-54.
20	517	6. Ombashi S, Kurniawan M, Koudstaal MJ, Allori AC, Jansson K, Rogers-Vizena CR, et al.
21	518	"Most efficient and meaningful patient reported appearance assessment in different cleft types
22 23	519	and age-groups with Cleft-Q". Plast Reconstr Surg. 2023.
24	520	7. Mossey PA, Little J, Munger RG, Dixon MJ, Shaw WC. Cleft lip and palate. Lancet.
25	521	2009;374(9703):1773-85.
26	522	8. Kummer AW. Types and causes of velopharyngeal dysfunction. Semin Speech Lang.
27	523	2011;32(2):150-8.
20	524	9. Knight J, Cassell CH, Meyer RE, Strauss RP. Academic outcomes of children with isolated
30	525	orofacial clefts compared with children without a major birth defect. Cleft Palate Craniofac J.
31	526	2015;52(3):259-68.
32	527	10. Van Lierde KM, Dhaeseleer E, Luyten A, Van De Woestijne K, Vermeersch H, Roche N.
33 34	528	Parent and child ratings of satisfaction with speech and facial appearance in Flemish pre-
35	529	pubescent boys and girls with unilateral cleft lip and palate. Int J Oral Maxillofac Surg.
36	530	2012;41(2):192-9.
37 20	531	11. Peterson-Falzone SJ. Speech outcomes in adolescents with cleft lip and palate. Cleft
30 39	532	Palate Craniofac J. 1995;32(2):125-8.
40	533	12. Baek RM, Kim BK, Jeong JH, Ahn T, Park M, Han J. The effect of age at surgery and
41	534	compensatory articulation on speech outcome in submucous cleft palate patients treated with
42	535	double-opposing Z-plasty: A 10-year experience. J Plast Reconstr Aesthet Surg. 2017;70(5):646-
43 44	536	52.
45	537	13. Markham C, van Laar D, Gibbard D, Dean T. Children with speech, language and
46	538	communication needs: their perceptions of their quality of life. Int J Lang Commun Disord.
47	539	2009;44(5):748-68.
48 40	540	14. Nicola K, Watter P. Health-related quality of life from the perspective of children with
50	541	severe specific language impairment. Health Qual Life Outcomes. 2015;13:127
51	542	15. Sell D, Sweeney T. Percent Consonant Correct as an Outcome Measure for Cleft Speech
52	543	in an Intervention Study. Folia Phoniatr Logop. 2019:1-9.
53 54	544	16. Klassen AF, Tsangaris E, Forrest CR, Wong KW, Pusic AL, Cano SJ, et al. Quality of life of
55	545	children treated for cleft lip and/or palate: a systematic review. J Plast Reconstr Aesthet Surg.
56	546	2012;65(5):547-57.
57		
58 59		24
~		

1 ว		
2	E17	17 Klasson AE Biff KM/M Longmiro NM Albert A Allen CC Avdin MA et al Develometric
4	547	findings and normative values for the CLEET O based on 2424 shildren and young adult nationts
5	546	with cloft lin and/or palato from 12 countries. CMAL 2019:100(15):E4EE E62
6 7	549	Wong Piff KW, Tsangaris E. Goodacro T. Forrost CP. Ducis AL. Cano SL et al. International
7 8	550	10. Wong Kill KW, Tsangalis E, Goodacie T, Forrest CK, Pusic AL, Carlo SJ, et al. International
9	551	multiphase mixed methods study protocol to develop a cross-cultural patient-reported
10	552	Once 2017;7(1):e0154C7
11	553	Open. 2017;7(1):e015467.
12	554	19. I Sangaris E, Wong Riff KWY, Goodacre T, Forrest CR, Dreise W, Sykes J, et al. Establishing
13	555	Content validity of the CLEFT-Q: A New Patient-reported Outcome Instrument for Cleft
15	550	Lip/Palate. Plast Reconstr Surg Glob Open. 2017;5(4):e1305.
16	557	20. Wong RITT KWY, Isangaris E, Goodacre TEE, Forrest CR, Lawson J, Pusic AL, et al. What
17	558	Matters to Patients with Cleft Lip and/or Palate: An International Qualitative Study Informing
18 10	559	the Development of the CLEFT-Q. Cleft Palate Craniofac J. 2018;55(3):442-50.
20	560	21. McLeod S, Crowe K, Shahaeian A. Intelligibility in Context Scale: Normative and
21	561	Validation Data for English-Speaking Preschoolers. Lang Speech Hear Serv Sch. 2015;46(3):266-
22	562	76.
23	563	22. McLeod S. Intelligibility in Context Scale: cross-linguistic use, validity, and reliability.
24 25	564	Speech, Language and Hearing. 2020;23(1):9-16.
25 26	565	23. McLeod S, Harrison LJ, McCormack J. The intelligibility in Context Scale: validity and
27	566	reliability of a subjective rating measure. J Speech Lang Hear Res. 2012;55(2):648-56.
28	567	24. McLeod S, Harrison LJ, McCormack J. Skattning av förståelighet i kontext: Svenska
29	568	[Intelligibility in Context Scale: Swedish] (T. Lagerberg, Trans.) Bathurst, NSW, Australia: Charles
30 21	569	Sturt University. 2012.
32	570	25. McLeod S, Harrison LJ, McCormack J. Schaal voor verstaanbaarheid in de context:
33	571	Nederlands [Intelligibility in Context Scale: Dutch] (A. van Doornik, Trans). Bathurst, NSW,
34	572	Australia: Charles Sturt University. 2013.
35	573	26. Lohmander A, Hagberg E, Persson C, Willadsen E, Lundeborg I, Davies J, et al. Validity of
36 27	574	auditory perceptual assessment of velopharyngeal function and dysfunction – the VPC-Sum and
38	575	the VPC-Rate. Clinical Linguistics & Phonetics. 2017;31(7-9):589-97.
39	576	27. The R Project for Statistical Computing: The R Foundation; 2021 [Available from:
40	577	https://www.r-project.org/.
41	578	28. Semb G, Enemark H, Friede H, Paulin G, Lilja J, Rautio J, et al. A Scandcleft randomised
42 43	579	trials of primary surgery for unilateral cleft lip and palate: 1. Planning and management. J Plast
44	580	Surg Hand Surg. 2017;51(1):2-13.
45	581	29. Long RE, Jr., Hathaway R, Daskalogiannakis J, Mercado A, Russell K, Cohen M, et al. The
46	582	Americleft study: an inter-center study of treatment outcomes for patients with unilateral cleft
47	583	lip and palate part 1. Principles and study design. Cleft Palate Craniofac J. 2011;48(3):239-43.
48 40	584	30. Shaw WC, Semb G, Nelson P, Brattstrom V, Molsted K, Prahl-Andersen B, et al. The
49 50	585	Eurocleft project 1996-2000: overview. J Craniomaxillofac Surg. 2001;29(3):131-40; discussion
51	586	41-2.
52	587	31. Lim CR, Harris K, Dawson J, Beard DJ, Fitzpatrick R, Price AJ. Floor and ceiling effects in
53	588	the OHS: an analysis of the NHS PROMs data set. BMJ open. 2015;5(7):e007765-e.
54 55		
56		
57		
58		25
59 60		For peer review only - http://hmiopen.hmi.com/site/about/quidelines.yhtml
00		

BMJ Open: first published as 10.1136/bmjopen-2023-071571 on 28 December 2023. Downloaded from http://bmjopen.bmj.com/ on June 14, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

BMJ Open

32. Stucki G, Liang MH, Stucki S, Katz JN, Lew RA. Application of statistical graphics to facilitate selection of health status measures for clinical practice and evaluative research. Clin Rheumatol. 1999;18(2):101-5. 33. Tsangaris E, Riff K, Vargas F, Aguilera MP, Alarcon MM, Cazalla AA, 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. Tsangaris E, Wong Riff KWY, Dreise M, Stiernman M, Kaur MN, Piplani B, et al. 34. Translation and cultural adaptation of the CLEFT-Q into Arabic, Dutch, Hindi, Swedish, and Turkish. European Journal of Plastic Surgery. 2018;41(6):643-52. Stiernman M, Klintö K, Persson M, Becker M. Comparison of Corresponding Scores From 35. the Cleft Hearing Appearance and Speech Questionnaire (CHASQ) and CLEFT-Q in Swedish Patients With Cleft Lip and/or Palate. The Cleft Palate-Craniofacial Journal. 2020:105566562096412. 36. Hammarström IL, Nyberg J, Alaluusua S, Rautio J, Neovius E, Berggren A, et al. Scandcleft Project Trial 2-Comparison of Speech Outcome in 1- and 2-Stage Palatal Closure in 5-Year-Olds With UCLP. Cleft Palate Craniofac J. 2020;57(4):458-69. 37. Lohmander A, Willadsen E, Persson C, Henningsson G, Bowden M, Hutters B. Methodology for speech assessment in the Scandcleft project--an international randomized clinical trial on palatal surgery: experiences from a pilot study. Cleft Palate Craniofac J. 2009;46(4):347-62. 38. Shprintzen RJ, Marrinan E. Velopharyngeal insufficiency: diagnosis and management. Curr Opin Otolaryngol Head Neck Surg. 2009;17(4):302-7. Brondsted K, Grunwell P, Henningsson G, Jansonius K, Karling J, Meijer M, et al. A 39. phonetic framework for the cross-linguistic analysis of cleft palate speech. Clinical Linguistics & Phonetics. 1994;8(2):109-25. 40. Grunwell P, Brondsted K, Henningsson G, Jansonius K, Karling J, Meijer M, et al. A six-centre international study of the outcome of treatment in patients with clefts of the lip and palate: the results of a cross-linguistic investigation of cleft palate speech. Scand J Plast Reconstr Surg Hand Surg. 2000;34(3):219-29. Sell D, Harding A, Grunwell P. A screening assessment of cleft palate speech (Great 41. Ormond Street Speech Assessment). Eur J Disord Commun. 1994;29(1):1-15. Sell D, Harding A, Grunwell P. GOS.SP.ASS.'98: an assessment for speech disorders 42. associated with cleft palate and/or velopharyngeal dysfunction (revised). Int J Lang Commun Disord. 1999;34(1):17-33. 43. John A, Sell D, Sweeney T, Harding-Bell A, Williams A. The cleft audit protocol for speech-augmented: A validated and reliable measure for auditing cleft speech. The Cleft palate-craniofacial journal : official publication of the American Cleft Palate-Craniofacial Association. 2006;43(3):272-88. Ahl R, Harding-Bell A. Comparing Methodologies in a Series of Speech Outcome Studies: 44. Challenges and Lessons Learned. The Cleft Palate-Craniofacial Journal. 2017;55(1):35-44. Chapman KL, Baylis A, Trost-Cardamone J, Cordero KN, Dixon A, Dobbelsteyn C, et al. 45. The Americleft Speech Project: A Training and Reliability Study. Cleft Palate Craniofac J. 2016;53(1):93-108. For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26	632 633 634 635 636 637 638 639 640 641 642 643 644	 46. Britton L, Albery L, Bowden M, Harding-Bell A, Phippen G, Sell D. A cross-sectional cohort study of speech in five-year-olds with cleft palate ± lip to support development of national audit standards: benchmarking speech standards in the United Kingdom. Cleft Palate Craniofac J. 2014;51(4):431-51. 47. Baylis A, Chapman K, Whitehill TL, The Americleft Speech G. Validity and Reliability of Visual Analog Scaling for Assessment of Hypernasality and Audible Nasal Emission in Children With Repaired Cleft Palate. Cleft Palate Craniofac J. 2015;52(6):660-70. 48. Pereira VJ, Tuomainen J, Lee KYS, Tong MCF, Sell DA. A perceptual outcome measure of velopharyngeal function based on the Cleft Audit Protocol for Speech-Augmented (CAPS-A VPC-Sum): Validation through a speech osteotomy study. Int J Lang Commun Disord. 2021;56(4):754-67.
27 28 29		
30 31		
32		
33 34 35		
35 36		
37 38		
39 40		
41 42		
43 44		
45 46		
47 48		
49		
50 51		
52 53		
54 55		
56 57		
58		27
59 60		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

2		
3 4 5 6	645 646 647	FIGURE LEGENDS
7 8	648	Figure 1. Correlations in patients with CP and CL(A)P
9 10 11	649 650	All correlations in both cleft types appeared significant (P<0.05), except for the correlation between the PCC and CLEFT-Q SDistress in patients with CL(A)P ($p = 0.285$). Note that VPC is inversely scored (higher
12	651	numbers correspond to worse outcomes) thus accounting for the negative correlations with the other
13 14	652	scales.
15 16 17	653	
17 18 19	654	
20 21	655	Figure 2: cross-sectional trend analyses of the age-groups
22 23	656	Analyses are presented per outcome measure, per cleft type.
24 25	657	
26 27	658	
28 29	659	Figure 3 Overview of the new proposed ICHOM Cleft Lip and Palate set concerning speech assessment
30 31	660	Newly made recommendations are coloured in pink
32 33	661	*Suggestion for centres that have adequate resources to implement and are interested in research with
34	662	speech outcomes.
35		
36	663	
37		
38 39	664	
40	665	
41	666	Annondices
42	000	Appendices
43	667	
44 45	668	Appendix 1. Speech outcomes of ICHOM Cleft Lip and Palate and field test data
46		• • • • • • • • • • • • • • • • • • • •
47	669	PROMs filled out by the patients are colored in purple, and grey if filled out by the parent(s). Clinical
48 49	670	outcome measures are colored in yellow.
50	671	*Data derived from the field test only in these cohorts, as there is no measurement moment included in
51 52	672	ICHOM Cleft Lip and palate in these age-groups.
55 55	673	
56 57	674	Appendix 2. differences of outcome scores based on income status of the country
58		28
59		For near review only - http://bmionen.hmi.com/site/about/quidalinas.yhtml
bυ		For peer review only intep://binjopen.binj.com/site/about/guidelines.kitim

1 2		
2 3 4 5	675 676	Univariate analyses. Outcome scores of patients deriving from countries with a high income status were taken as reference. *Significant difference
6 7	677	
8 9	678	Appendix 3. descriptives for subgroup analyses
10 11	679	Descriptives of patients deriving from high income countries, including all patients from the ICHOM
12 13	680	centers.
14 15	681	
16 17	682	Appendix 4: maximum scores and VPC overview
17 18 19 20 21 22 23 24 25 26 27 28 90 31 32 33 34 35 36 37 38 90 41 42 43 45 46 47 48 90 51 52 54 55 56		
57 58		29
59 60		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml



60





All correlations in both cleft types appeared significant (P<0.05), except for the correlation between the PCC and CLEFT-Q SDistress in patients with CL(A)P (p = 0.285). Note that VPC is inversely scored (higher numbers correspond to worse outcomes), thus accounting for the negative correlations with the other scales.

270x119mm (300 x 300 DPI)




Page 35	of 55			BMJ Open ft, inc 77		
1	5-7 years	8-9 years	10-13 years	14-16 years 🛓 🖁	17-19 years	20-22 years
2				for us		
3 4 5	ICS	SFunction *	ICS	SFunction * eiger	SFunction *	PCC
6 7 8	PCC	SDistress *	PCC	SDistress * to support	SDistress *	VPC
9 10 11 12	VPC		VPC	nloaded fr perieur (A t and data		SFunction
13 14 15 16			SFunction	om http://t BES) · mining, Al		SDistress
17 18 19			SDistress	training,		
20 21 22				mj.com/ c and simil		
25 24 25 26				n June 1. ar techno		
27 28 29				4, 2025 at logies.		
30 31 32				Agence		
33 34 35				Bibliogra		
36 37 38			For peer review only - http://bu	niopen hmi com/site/about/quidelines	xhtml	
39 40			For peer review only intep.//bi		Artarin	

Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

BMJ Open

	Income	Beta	SE	P-value	Difference between groups	P-value
ed	High (ref)	72.10	0.51		High VS Middle	< 0.001*
elat s	Middle	-16.39	1.58	<0.001*		
-Re						
ing	High (ref)	72.10	0.51		High VS Low	<0.001*
eak	Low	-13.08	1.68	<0.001*		
Spe						
					Middle VS Low	0.285
	High (ref)	70.87	0.52		High VS Middle	<0.001*
r) ion	Middle	-12.32	1.64	<0.001*		
eec						
Sp Fu	High (ref)	70.87	0.52		High VS Low	<0.001*
	Low	-13.18	1.71	<0.001*		
		Z			Middle VS Low	0.923

Appendix 1: differences of outcome scores based on income status of the country

Univariate analyses. Outcome scores of patients deriving from countries with a high income status were taken as reference. *Significant difference

2	
3	
1	
-	
5	
6	
7	
Q	
0	
9	
1	0
1	1
1	n
1	2
1	3
1	4
1	5
1	6
1	-
1	/
1	8
1	9
ר	ñ
2	0
2	1
2	2
2	3
2	1
2	4
2	5
2	6
2	7
2	, 0
2	0
2	9
3	0
З	1
2	ר
2	2
3	3
3	4
3	5
2	6
2	-
3	/
3	8
3	9
1	ñ
4	1
4	I.
4	2
4	3
۵	Δ
7	-
4	5
4	6
4	7
4	8
,	0
4	9
5	0
5	1
5	2
5	2
2	د
5	4
5	5
5	6
5	7
ر -	′ ~
-5	б

59 60 Missing

	CP (n=762)	CL(A)P (n=1379)	Total (n=2141)
Gender			
Male	308 (40.4%)	904 (65.6%)	1212 (56.6%)
Female	454 (59.6%)	475 (34.4%)	929 (43.4%)
SDistress			
Number	517 (67.8%	1068 (77.4%)	1585 (74.0%)
Missing	245 (32.2%)	311 (22.6%)	556 (26.0%)
SFunction			
Number	519 (68.1%)	1076 (78.0%)	1595 (74.5%)
Missing	243 (31.9%)	303 (22.0%)	546 (25.5%)
VPC		0	
Number	173 (22.7%)	213 (15.4%)	386 (18.0%)
Missing	589 (77.3%)	1166 (84.6%)	1755 (82.0%)
ICS			
Number	169 (22.2%)	228 (16.5%)	397 (18.5%)
Missing	593 (77.8%)	1151 (83.5%)	1744 (81.5%)
PCC score			
Number	152 (19.9%)	192 (13.92%)	344 (16.1%)

610 (80.1%)

Appendix 2: descriptives for subgroup analyses

Descriptives of patients deriving from high income countries, including all patients from the ICHOM centers.

1187 (86.1%)

1797 (83.9%)

Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

Appendix 3: maximum scores and VPC overview

		Number of measurments with max score	Total measurements
СР	SDistress	18.8% (97)	517
	SFunction	19.8% (103)	519
	ICS	29.0% (49)	169
CL(A)P	SDistress	15.2% (162)	1068
	SFunction	15.9% (172)	1076
	ICS	18.0% (41)	228
Maximum sco	res		

		СР		CL(A)P			
	Competent	Marginally	Incompetent	Competent	Marginally	Incompetent	
	(0)	incompetent	(2)	(0)	incompetent	(2)	
		(1)			(1)		
5 years	43 (50%)	21 (24.4%) 🧹	22 (25.6%)	11 (15.5%)	17 (23.9%)	43 (60.6%)	
12 years	39 (65%)	19 (31.7%)	2 (3.3%)	46 (52.3%)	28 (31.8%)	14 (15.9%)	
22 years	11 (40.7%)	13 (48.1%)	3 (11.1%)	26 (48.1%)	19 (35.2%)	9 (16.7%)	
vi e overview							

9 of 55		BMJ Open	open-20 oy copy	}	
STROBE Stateme	ent—che	cklist of items that should be included in reports of observational studies	23-071571 , rright, inclu		
	Item No.	Recommendation	on 28 D ding foi	Page No.	Relevant text from manuscript
Title and abstract	1	(<i>a</i>) Indicate the study's design with a commonly used term in the title or the abstract	cember 2023. Downloade ∧Enseignement Superieu uses related to text and d		An international, multicenter study was set up including centers from the United State and The Netherlands. Outcon of clinical measures and Pati Reported Outcome Measures(PROMs) were
			r (AE ata r		according to the ICHOM set
		(b) Provide in the abstract an informative and balanced summary of what was done and what w found	achttp://bmjopen.bmj.co ES) ining, Al training, and s		Data from 2500 patients was included. Measured outcome contained univariate regress analyses, trend analyses, T-t correlations and floor and ceiling effects.
			m/ on June 14, 2025 at imilar technologies.		Results PROMs correlated low to moderate with clinical outco measures. Clinical outcome measures correlated low to moderate with each other to
			Agence Bibliograph		contrast, two CLEFT-Q scal- correlated strongly with each other. All PROMs and the Percent Consonants Correct(PCC) showed an eff of age. In patients with an
		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xł	ntml de		

	BMJ Open Copy Y	Pa
	3-071571 on 28 Dec I I I I I I I I I I I I I I I I I I I	isolated cleft palate, a ceiling effect was found in the intelligibility in context scale(ICS).
Introduction		
Background/rationale 2	Explain the scientific background and rationale for the investigation being reported to text and data mining. At training, and similar technologies.	Recently, the International Consortium for Health Outcomes Measurement (ICHOM) developed the ICHOM Standard Set for Cleft Lip and Palate (ICHOM Standard Set), with different pathways for varying cleft types(5). Based on patient and expert consensus, a minimal, accessible set of outcome measures was established to enable benchmarking between cleft centers in a systematic manner. For speech assessment an outcome set was included with both clinical measures and Patient Reported Outcome Measure (PROMs), being the patient's and parent's perspectives. So far, the selected standardized speech outcome measures and
	ibliogr	their timing have not been evaluated. As an increasing
	For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	

41 of 55	BMJ Open	open-20 by cop
	For peer review.	 Nincluding for uses related to text and optimize this ICHOM Boston Children's Hospital (Boston, USA), Duke University Hospital (Durham, USA), and the Erasmus Medical Center (Rotterdam, The Netherlands), started clinical implementation and an international collaboration in 2015. The overarching aim of this collaboration is to share data and knowledge obtained by using the set in standard care. Additionally, they collaborate with McMaster University (Hamilton, Canada), who developed and field tested the CLEFT-Q questionnaire, of which many scales are included in the ICHOM Standard set.
Objectives	3 State specific objectives, including any prespecified hypotheses	The objective of this study was to evaluate the current standardized speech outcome measures of the ICHOM Standard Set for patients with CP±L. More specifically, the value of every speech outcome measure was examined as well
	For peer review only - http://bmjopen.bmj.com/site/about/gu	idelines.xhtml

jopen-2023-071571 on 28 December 2023. Downloaded from http://bmjopen.bmj.com/ on June 14, 2025 at Agence BibfFographique de I Enseignement Superieur-(ABES) . J by copyright, including for uses related to text and data mining, Al training, and similar technologies. BMJ Open as the best age intervals for assessment of these outcome measures. In addition, other speech assessment tools are discussed. Finally, recommendations are made for an optimal and complete 10 assessment of speech in patients 11 12 with CP±L, that is efficient and 13 accessible for all cleft centers. 14 . une paper Methods 15 16 Present key elements of study design early in the paper Study design 4 Three centers (Boston 17 Children's Hospital, Duke 18 University Hospital, and 19 Erasmus Medical Center,) each 20 implemented the ICHOM 22 Standard Set in 2015. All 23 patients treated at these centers 24 for a cleft palate with a cleft 25 26 lip/cleft alveolus (CL(A)P), or 27 an isolated cleft palate (CP) who 28 were assessed according to the 29 ICHOM Standard Set (age 30 31 range 5-22 years), were 32 included. In addition, another 33 patient group derived from an 34 international field test of the 35 36 CLEFT-Q, by McMaster 37 University(17). 38 Setting 5 Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, Data was collected 39 follow-up, and data collection restrospectively over a 6 year 40 41 42 43 For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml 44 45 46

1 2

3

4

5

6 7

8

9

Page 43 of 5	55	BMJ Open	ijopen-2023-0 J by copyrigh
1	Participants 6	(a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case control study—Give the eligibility criteria, and the sources and methods of ease ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants	-071571 on 28 December 2023. Downloaded from http://bmjopen.bmj.com/ on June 14, 2025 at Agence Bibliogu Enseignement Superieur (ABES). pht, including for uses related to text and data mining, Al training, and similar technologies.
41 42 43 44		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtr	nl de l

44 45 46 period (2015-2020) and extracted from the electronic patient files in 2018 and 2020

According to the ICHOM protocol, both CLEFT-Q Speech scales were assessed at ages 12 and 22 years(fig. 1). Both PCC and VPC were scored at ages 5, 12 and 22 years, and ICS at ages 5 and 12 years. The field test cross-sectionally collected data from patients with a cleft across 12 different countries with different income-

(as a data update).

statuses(17).

All patients treated at these centers for a cleft palate with a

(CL(A)P), or an isolated cleft palate (CP) who were assessed according to the ICHOM Standard Set (age range 5-22 years), were included. In addition, another patient group derived from an international field test of the CLEFT-Q, by McMaster University(17). According to the age cut-off of the ICHOM Standard Set, only outcomes from CLEFT-Q scales

cleft lip/cleft alveolus

(...)

	BMJ Open Sy Spen-202	Pag
	(b) Cohort study For matched studies, give matching criteria and number of exposed and unexposed Case-control study For matched studies, give matching criteria and the number of controls per case	of field test patients with a CP±L up to 22 years old were included in the current study(fig 1). Patients from the participating centers were excluded in case they could not sufficiently speak or write the language native to the center's country .
Variables	 Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. At training, and similar technologies. 	To prevent possible influence of income-status on the outcomes, univariate regression analyses were used to examine differences in outcome scores of the SDistress and SFunction before further analyses. Data was categorized according to the income status of the country where the data had been collected. In order to examine the added value of each PROM and clinical outcome measure in regard with the other measures, correlations were examined between the PROMs; between the clinical outcome measures; and between the PROMs and
	For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	

For peer review on	71571 on 28 December 2023. Downloaded from http://bmjopen.bmj.com/ on June 14, 2025 at Agence B Enseignement Superieur (ABES) . t, including for uses related to text and data mining, Al training, and similar technologies.	the clinical outcome measures. Pearson correlations were used, and outcomes were analysed per cleft type. Correlations were considered strong in case r>0.7; moderate between r=0.5 and R= 0.7; and weak in case r<0.5. Analyses within and between different age-groups were done to explore whether the current outcome measures are assessed at the optimal age-points, and whether additional measurement moments are indicated either for PROMs or clinical outcome measures. Therefore, not only time-points of the ICHOM protocol were included in analyses. As CLEFT-Q outcome scores of all ages between 8-22 years were included from the field test data, time-points used by other large initiatives as Eurocleft, Scandcleft and Americleft were considered as well (28-30). Doing so, the following age-groups were set up: 5-7 years; 8-9 years; 10-13 years; 14-16 years; 17-19 years,
	Agence Bibliographi	following age-groups were set up: 5-7 years; 8-9 years; 10-13 years; 14-16 years; 17-19 years, 20-22 years(fig. 1). Per age-group, possible differences between scale scores

jopen-2023-071571 on 28 December 2023. Downloaded from http://bmjopen.bmj.com/ on June 14, 2025 at Agence Bibliograph Enseignement Superieur (ABES) . d by copyright, including for uses related to text and data mining, Al training, and similar technologies.

ique de l

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1 2

3

4

5

6

7

8

9

10

11 12

13

14

15

16 17

18

19

20

21 22

23

24

25

26 27

28

29

30

31 32

33

34

35 36

37

38

39

44 45 46

jopen-2023-071571 on 28 December 2023. Downloaded from http://bmjopen.bmj.com/ on June 14, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES) . by copyright, including for uses related to text and data mining, Al training, and similar technologies. were examined with independent T-tests. Bonferroni correction was applied for rtraining, and sin. multiple testing. Trend analyses were performed to identify potential problems in specific age-groups. Floor- and ceiling effects were examined to identify the suitability of the outcome measures in our population. A floor or ceiling effect is seen when a considerable amount of the outcome scores are either scored the best (in this case a maximum score, thus a ceiling effect), or the worse (in this case a minimum score, thus a floor effect). Both effects result in a truncated distribution of the outcomes on either side of the scale (31, 32). Minimum and maximum score outcomes of all PROMs and the PCC were evaluated. A percentage of 20% or more of the patients scoring the minimum or maximum outcome score was considered as a ceiling effect. In VPC, the outcome distributions were examined. For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

e 47 of 55	BMJ Open BWJ Open 2023	
Data sources/ measurement	8* For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group 8. For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group 9. Compared to text and data multiply of assessment methods if there is more than one group	Outcome measuresPROMsCLEFT-Q ScalesThe CLEFT-Q is developedspecifically to assess QoL fromthe patient's perspective inpatients with a CP±L. Aliterature review, patientinterviews and psychometrictesting, established the finalcontent of the scales, whichcovers several overarchingdomains(18-20). Speech isassessed through two scales,each covering a differentdomain. Both scales have 3response options for each item(always; sometimes; never); alower score equals a worseoutcome. Completing the scalescan be done online; it will takethe patient several minutes.Speaking-Related Distress(SDistress) is part of thepsychosocial domain. The scalecontains 10 items that relate tothe psychosocial part of speechdifficulties, like nervousness orfrustration(20).Speech Function (SFunction)focuses on the functional speechdifficulties that patients
	For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	

1 2

3

4

5 6

7

8

9 10

11 12

13

14

15

16 17

18

19

20

21 22

23

24

25

26 27

28

29

30

31 32

33

34

35

36 37

38

39 40 41

42

jopen-2023-071571 on 28 December 2023. Downloaded from http://bmjopen.bmj.com/ on June 14, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES) . by copyright, including for uses related to text and data mining, Al training, and similar technologies. themselves identify, for example the ability to say certain letters or words. The scale consists of 12 items that For beer review on. belong to the facial function domain(20). Intelligibility in Context Scale(ICS) is a measure that assess the intelligibility of the child. It is a 7-item, parentreported questionnaire designed to be scored by speech pathologists. The score indicates a child's level of functional intelligibility, by assessing the degree to which the speech of the patient is understood by different communication partners. The total score is calculated by the averages of the items completed. ICS appeared to be a valid and reliable tool for children with speech disorders, (21, 22), but not specifically designed nor validated for patients with $CP\pm L(23)$. It is available in several languages, and normative data exists for English-speakers(24, 25). Clinical outcome measures

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

* 71 57 inc	Percent Consonants
Sud o	Correct(PCC) is developed to
ling 2	detect speech sound errors. PCC
for Do	scores are calculated by using a
	standard, crossecitonally
est a series and a s	translated set of words that
	include all speech problems
ed t	children with CP±L often tend
	to have.
	In case of any problems, their
ande	severity can be categorized:
dated	PCC scores of 85-100% indicate
	mild to no problems; scores of
	65-84.9% indicate mild-
â. b	moderate problems; scores of
	50-64.9% indicate moderate-
ani și	severe problems; and scores
, pg , pg	<50% indicate severe
an <u>a</u>	problems(20). PCC is suitable
si g	for usage in patients with CP±L
	when assessed by well-trained
r tech	clinicians(8).
ie 14, inolo	Velopharyngeal Competence
gier gier gier gier gier gier gier gier	rating(VPC) discriminates
en e	between three categories:
	'competent', 'marginally
	incompetent' and 'incompetent'.
	The outcome is determined by
	the speech therapist based on

the PCC test and spontaneous

speech. In case of any clinical

jopen-2023-071571 on 28 December 2023. Downloaded from http://bmjopen.bmj.com/ on June 14, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES) . by copyright, including for uses related to text and data mining, Al training, and similar technologies. For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

		BMJ Open	י הלאלי האלי	jopen-2023	Page 5
			jii, including to uses related to text and nata mining	-071571 on 28 December 2023. Downloaded from http: Enseignement Superieur (ABES) .	evidence of minor problems regarding the competence, VPC was categorized as 'marginally incompetent'. When clinically significant problems were detected, suggesting surgical management and/or speech therapy, VPC was categorized as 'incompetent'. Prior studies found VPC to be suitable as a first clinical choice for the assessment of velopharyngeal dysfunction and is recommended for both clinical follow-up and research(26)
Bias	9	Describe any efforts to address potential sources of bias			tonow-up and research(20).
Study size	10	Explain how the study size was arrived at		NA (see data conjection, which is cited at point 7 "participants")	
Continued on next page		For peer review only - http://bmjopen.bmj.co	rom/site/about/guidelines.xhtml	n June 14, 2025 at Agence Bibliographique de	

55			BMJ Open Sy pen-2023	
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	See including to 28 December 29 December 2	
Statistical methods	12	 (a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions 	See citations from point 7 'variables' See citations from point 7 (vitil b) 5 (vitil b) 5	
		(c) Explain how missing data were addressed(d) Cohort study—If applicable, explain howloss to follow-up was addressedCase-control study—If applicable, explain howmatching of cases and controls was addressedCross-sectional study—If applicable, describeanalytical methods taking account of samplingstrategy(e) Describe any sensitivity analyses	Variables N/A See citations from point 7 'variables' N/A	
Results			gies	
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	N/A at Agence Biblio	
		For peer review only -	http://bmjopen.bmj.com/site/about/guidelines.xhtml	

					2023- pyrigl		
		(b) Give reasons for non-participation at each stage	N/A, as our data was collected as part of regular care protocols		071571 on 28 December 2 Enseign ht, including for uses rela		
		(c) Consider use of a flow diagram	•	_	023. ted t	_	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information	12, table 1		o n D Fieldရန္အရပ္ က = သူမ်ိဳးရင္	ICHOM (n = 777)	Overall (n = 2500)
		on exposures and potential confounders		Sex	ieur d da		
				Male	555 BP20	444 (57.1%)	1425 (57.0%)
				Female	742 (3 .1	333 (42.9%)	1075 (43.0%)
				Cleft Type	Al tr		
				СР	517 (35).0%	301 (38.7%)	818 (32.7%)
				CL(A)P	1206 (20.0 2)	476 (61.3%)	1682 (67.3%)
				Income classifications	l simi		
				High Income	1364 (Z9.2%)	777 (100%)	2141 (85.6%)
				Up Middle income		0 (0%)	199 (8.0%)
				Low Middle Income	160 (92,3%)	0 (0%)	160 (6.4%)
		(b) Indicate number of participants with missing		N/A	<u>s.</u>		
		data for each variable of interest			> 		
		(c) Cohort study — Summarise follow-up time			nce		
Outcome	15*	Cohort study — Report numbers of outcome			Bibl		
-	-	events or summary measures over time			iogr		

je 53 of 55		BMJ Open by copyrig ig
	<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	-071571 on 2
	<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	See point 14 'descriptive data', where table bis cited.
Main results 16	 (a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and 	N/A es related to tex
	(<i>b</i>) Report category boundaries when continuous	Page 10: Psychometric validation of the Barrier and SFunction confirmed suitability to
	variables were categorized	a 0 to 100 scale deriving from the sum score for analysis(17).
	(<i>c</i>) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Not relevant for the current study
Continued on next page		pen.bmj.com/ on June 14, 2025 at Agence Bibliog ning, and similar technologies.
	For peer review only - http://b	njopen.bmj.com/site/about/guidelines.xhtml

3 4

	BMJ Open	d by copyrig 23.
Other analyses	17 Report other analyses done—eg analyses of subgroups and interactions, and sensitivity a	InalysesAll participating ICHOM centers are high-income countries, whereas part of the field test data was collected in upper middle and
Discussion	(0)	
Key results	18 Summarise key results with reference to study objectives	From the current study, it can be concluded that the current ICHOM Standard Set is informative and efficient. PROMs were shown to be of added value, and the CLEFT- Q appeared to be the most suitable PROM. Therefore, continuation of collecting the current outcome measures and time points is recommended. Furthermore, a minor extension is suggested: In addition to the current time-points of assessment, it is recommended to implement the CLEFT-Q SDistress scale at the age of 8-9
	For peer review only - http://bmjopen.bmj.com/site/about	/guidelines.xhtml de

ge 55 of 55		BMJ Open	open-20; by copy	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	23-071571 on 28 December 2023. Downloaded from http://bmjopen.bm Enseignet@ent Superieur (ABES) . rright, including for uses related to text and data mining, Al training, ar	and 17-19 as well. Further adjustments of the set could comprise an additional categorization of the PCC score, based on the framework of Eurocleft and adjusted for clinica usage. Data was analyzed cross-sectiona Longitudinal analyses to explore development of speech and for benchmarking will be possible in the future, since data collection continues. Moreover, because thi study included data from the CLEFT-Q field test, a higher number of outcome data from the CLEFT-Q scales was available for analyses than from the other outcome measures included in the
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	si. iiii 1520 (entire technologies) technologies techn	ICHOM Standard Set.
		17 For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xh	tml de l	

3 4

	BMJ Open	jopen-2023 I by copyrig	Page
		Jh -07 incomide incom	erations ons are
Generalisability 21	Discuss the generalisability (external validity) of the study results	ecember 2023. Downloaded from http://bmjopen.bmj.com/ on June 14, 2025 at . 定年nseignement Superieur (ABES) . r uses related to text and data mining, Al training, and similar technologies.	From the current study, it can be concluded that the current ICHOM Standard Set is informative and efficient. PROMs were shown to be of added value, and the CLEFT- Q appeared to be the most suitable PROM. Therefore, continuation of collecting the current outcome measures and time points is recommended. Furthermore, a minor extension is suggested: In addition to the current time-points of assessment, it is recommended to implement the CLEFT-Q SDistress scale at the age of 8-9 and 17-19 as well. Further adjustments of the set could comprise an additional categorization of the PCC score, based on the framework of Eurocleft and adjusted for clinical usage.
Other information		Age	
Funding 22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	23ce Bibliograg	This project received no funding. However, S. Ombashi works as a PhD-student for an European Union-funded Network, the
	18 For peer review only - http://bmjopen.bmj.com/site/about/guidelines.	chtml de l	

Page 57 o	7 of 55 BMJ Open Copy 202	
1 2 3 4 5 6 7 8 9 10 11	3-071571 on 28 December 202: Enseignem ight, including for uses related	'Eureopean Reference Network for Craniofacial anomalies and Ear, Nose and Throat disorders'. The results and findings from this project are of help in further European alignment concerning standardized outcome measures in cleft care
12 13	*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in controls in case-control studies and if applicable for exposed and unexposed groups in controls in case-control studies and if applicable for exposed and unexposed groups in controls in case-control studies and if applicable for exposed and unexposed groups in controls in case-control studies and if applicable for exposed and unexposed groups in controls in case-control studies and if applicable for exposed and unexposed groups in controls in case-control studies and if applicable for exposed and unexposed groups in controls in case-control studies and if applicable for exposed and unexposed groups in controls in case-control studies and if applicable for exposed and unexposed groups in controls in case-control studies and if applicable for exposed and unexposed groups in controls in case-control studies and if applicable for exposed and unexposed groups in controls in case-control studies and if applicable for exposed and unexposed groups in controls in case-control studies and if applicable for exposed and unexposed groups in controls in case-control studies and its controls in case-control studies and con	l cross-sectional studies.
15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32	Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicines.std http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.g. Al training, and similar technologie and similar technologie at the state of the	transparent reporting. The STROBE nnals of Internal Medicine at atement.org.
33 34 35 36 37 38 39 40 41	s. s.	
42 43 44 45 46	For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	