Published in final edited form as:
Cleft Palate Craniofac J. 2006 July ; 43(4): 383–391.

Health Professionals' Assessment of Health-Related Quality of Life Values for Oral Clefting by Age Using a Visual Analogue Scale Method

George L. Wehby, M.P.H.
College of Public Health, Department of Health Management and Policy, University of Iowa, Iowa City, Iowa.

Robert L. Ohsfeldt, Ph.D. [Professor]
Department of Health Policy and Management, School of Rural Public Health, Texas A&M University, College Station, Texas.

Jeffrey C. Murray, M.D. [Professor]
Department of Pediatrics, College of Medicine, University of Iowa, Iowa City, Iowa.

Abstract

Objective—To elicit health-related quality of life (HRQL) values associated with oral clefting by age using a visual analogue scale, and to explore the appropriateness of using health professionals as evaluators.

Methods—A representative group of health professionals working on craniofacial and/or cleft palate teams in the United States was sampled. Values (between 0 and 1) representing the HRQL associated with isolated and nonisolated oral clefting for infants, children, adolescents, and adults were obtained. The relationships between selected evaluator characteristics and values were also assessed.

Results—Of 330 professionals surveyed, 133 (40%) completed and returned reliable evaluations. Overall, HRQL values were clustered toward the right tail of the scale, indicating modest decreases in HRQL. Most evaluators reported feeling confident in completing the evaluations. HRQL values seemed to vary by team type (cleft palate only versus cleft palate/craniofacial care) and geographic location, but no major differences were found overall for any selected evaluator characteristics.

Conclusions—This study provides HRQL values for oral clefting based on preferences of health professionals that may be useful in evaluating the effectiveness and cost-effectiveness of prevention and treatment strategies, including those carried out in clinical trial studies. The clustered pattern of HRQL values suggests either a consensus among evaluators of a limited burden of oral clefting or an overall lack of understanding of the evaluation task.

Keywords

cleft lip; cleft palate; cost effectiveness; cost utility; health-related quality of life; health utilities; health values; oral clefting; quality-adjusted life years; visual analogue scale

Health-related quality of life (HRQL) is the term used to describe the impact of health status on the quality of human life. Although the phrase is often applied to a variety of definitions

Address correspondence to: Dr. Jeffrey C. Murray, Department of Pediatrics, College of Medicine, 2182 ML, University of Iowa, Iowa City, IA 52242-1181. E-mail jeff-murray@uiowa.edu..

This work was supported by NIH grants DE08559 and DE16215.

and specifications and for different purposes, HRQL reflects a perception of the undesirability of health or disease conditions or states in terms of their impact on quality of life. A "comprehensive" HRQL measure captures the impact of disease on physical functioning and comfort (pain), mental health, and social interactions. When combined with the quantity of life, HRQL provides a more complete indication of the impact of disease and treatments than either measure alone (e.g., quality-adjusted life years [QALYs]).

Assessment of HRQL within this framework involves the identification of preferences or values that describe the undesirability of health conditions or states compared to reference states. The HRQL values would be quantified through scores that could be combined with duration of life (e.g., QALYs) and used in conducting economic and effectiveness evaluations of health care interventions, including cost-effectiveness and cost-utility analyses, to guide the allocation of health care resources to improve health outcomes overall (Eisenberg, 1989). Further, because evidence-based medicine and the role of randomized clinical trials become increasingly important in clinical practice, it is critical to have reliable metrics to assess outcomes.

The HRQL scores used to adjust the duration of life for HRQL are usually specified on a continuum that ranges from low to high quality of life—specifically from death to perfect health (Gold et al., 1996). Scores of 0 and 1 would indicate death and perfect health states, respectively. The lower level of the scale can be extended beyond 0 to accommodate health states that are thought to be of lower value than death. Valuing a certain health state on this continuum involves assigning a score between 0 (or the lower value used) and 1 to indicate the relative desirability of this health state to the reference health states of death and perfect health that bound the continuum.

Several methods have been established and used to obtain HRQL scores. The most commonly used methods are the standard gamble (SG), the time-tradeoff (TTO), and the visual analogue scale (VAS) (Gold et al., 1996). The choice of which evaluation method(s) to use relates to both theoretical and practical considerations, such as ease of administration, cognitive burden imposed on evaluators, and time to complete the evaluation (Torrance, 1986). The SG method requires evaluators to choose between a gamble of death and perfect health on one side and a definite occurrence of the evaluated health state on the other, and thus identifies the probability of perfect health occurrence in the gamble (which corresponds to the HRQL score) at which the evaluators are indifferent between taking the gamble and the certain choice. This method has the advantage of being directly linked to the axioms of expected utility (Torrance and Feeny, 1989), but it carries a heavy cognitive burden for those rating health states. The TTO method is somewhat more intuitive than SG, which reduces cognitive burden. Evaluators report the amount of time they would be willing to give away to be in perfect health compared to continued living in the evaluated health state, with the HRQL score being the proportion of time at which evaluators are indifferent between remaining in the evaluated state or living in perfect health.

The most straightforward evaluation techniques involve direct rating and state comparison methods, including VAS. In simple terms, the task in VAS is simply to identify the location of a health state on a scale that ranges between the two selected reference states. However, there are several variations of VAS. The scale may be presented horizontally or vertically, the scale may be calibrated with "tick" marks or left unmarked between the two reference states, or additional reference health states may be marked on the scale. VAS has been used widely in various preference evaluation settings, primarily because VAS methods are simple and relatively easy to administer (especially to self-administer). Overall, VAS and utility-based methods (SG and TTO) have produced relatively different HRQL values, with VAS providing lower scores (Read et al., 1984; Torrance et al., 2001). The interrater and intrarater reliability

of direct rating methods has been reported in several studies to be relatively high (about 0.8) (Froberg and Kane, 1989; Krabbe et al., 1997).

In this study, we chose to obtain HRQL values related to oral clefting based on preferences of health professionals who serve on cleft palate and craniofacial teams in the United States. Oral clefting is one of the most common birth defects, with an incidence that varies by population and that ranges between 0.3 and 3.6 per 1000 live births (Mossey and Little, 2002). Oral clefting involves both isolated forms and nonisolated forms that involve the presence of other malformations and that comprise about 30% of clefting cases (Jones, 1988).

The impacts of oral clefting on individual health are substantial and have variable magnitude and scope by age as well as by type of clefting. During infancy and childhood, various health interventions are needed to target clefting and associated health complications, including surgical, medical, speech, dental, and other health interventions. Clefting is also associated with various health complications, including feeding, speech, growth, and physical health problems such as recurrent ear infections (Nackashi et al., 2002).

Several studies have described the psychological and social burden related to clefting through age. Children born with clefts have been reported to face a challenging psychological adjustment, mainly thought to result from a low level of satisfaction with facial appearance as well as inadequate acquisition of social skills to enable this adjustment (Kapp-Simon, 1986). In addition, a higher risk for developmental problems, including cognitive performance, has been reported (Jocelyn et al., 1996; Jelliffe-Pawlowski et al., 2003). Part of the treatment for clefting extends throughout adolescence, during which psychological adjustment has been reported to be a problem as well (Kapp-Simon et al., 1992; Thomas et al., 1997). Lower psychological adjustment, cognitive performance, as well as social and economic achievement in processes such as marriage, educational attainment, and income level, have been reported as well among adults with oral clefts in several studies (Heller et al., 1981; Bjornsson and Agustsdottir, 1987; Ramstad et al., 1995a, 1995b; Marcusson et al., 2001; Nopolous et al., 2002). No studies have reported the impact of clefting on the quality of life of infants. Survival in association with clefting has also been assessed, and recent work suggests that there may be an overall increase in mortality from all causes, with suicide as a single significant contributor; this is consistent with concerns about psychological adjustment (Christensen et al., 2004).

These findings indicate that there may be a relative burden of oral clefting on the overall quality of life across all age groups. However, no previous studies have attempted to directly measure preferences related to oral clefting as a health state and to provide HRQL scores that could be used in evaluating the effectiveness and cost effectiveness of health interventions and prevention programs aimed at clefting. Furthermore, obtaining HRQL scores would provide another tool to quantify the burden related to oral clefting.

The main objective of this study was to obtain and make available a set of HRQL values by age and clefting status using the VAS method. We also aimed to explore the appropriateness of using VAS-elicited preferences of health professionals involved in the care of individuals with clefts to measure the undesirability of clefting as a health state.

Methods

Setting and Participant Sample

We chose to obtain preferences of health care professionals rather than other groups, such as patients or parents, in this study. The main rationale was the exploratory nature of this work and concerns about the appropriateness or utility of such assessments among patients with oral clefts or their parents. Obtaining the preferences of professionals and experts who are highly

involved in evaluation of and provision of care to individuals with oral clefts from the early stages of life and through later stages seemed to be a reasonable approach to initiate measures of HRQL values related to oral clefting, particularly with respect to practical considerations. Health professionals are expected to be familiar with health and quality-of-life limitations and thus should require less effort to solicit their preferences. They may also provide a more informed and reliable evaluation compared to other groups.

A convenient sample of 330 members of cleft palate and craniofacial teams located in the United States was chosen. The sampling frame of the teams was based on the 2003–2004 membership team directory of the American Cleft Palate–Craniofacial Association (ACPA), which included a listing of teams (and their ACPA members) that responded to ACPA for inclusion in the directory. As a strategy to enhance response rate, the sample consisted mainly of listed team members who also attended the 61st annual meeting of the ACPA in March, 2004.

While the ACPA team directory may not provide a complete listing of all teams and members in the United States, it is currently the major available source of team listings. All team characteristic data used in this study, including specialties of members, geographic locations, and type of team (cleft and/or craniofacial care team), were based on information reported by the teams to ACPA and were obtained from this directory. Because of potential inaccuracies in this information, the implications of any comparisons of the HRQL scores involving those characteristics (such as differences in scores by specialty) should be carefully considered.

Study Survey

A draft survey instrument was pilot-tested among a group of 36 craniofacial, pediatric, and/or genetic experts at the University of Iowa who had at least some role in caring for patients with clefts. The purpose of the pilot test was to assess response rates, receive feedback on the clarity of the survey content, and to estimate the time required to complete the HRQL evaluation. Twenty-one experts (58%) responded to the survey, including 16 (44%) who completed the survey. The average time reported for completing the survey was 9 minutes. The pilot version of the survey included an example VAS with two health states (acute otitis media and advanced stage of breast cancer) that had been previously evaluated by professionals with a VAS scale method. Several respondents expressed concerns about the example scale influencing the evaluators' decision for clefting. Some concerns were also reported about the length and complexity of the instructions. Therefore, the survey was revised by taking out the example scale and by shortening and simplifying the instructions.

A two-page hardcopy of the final survey was sent to the selected sample of team members. Instructions were given about how to complete a HRQL evaluation on a VAS. Recipients were also asked to report the level of confidence in completing the HRQL evaluation by age group. The final survey instrument and the attached cover letter that were used are available on the following website: http://genetics.uiowa.edu/publications.html.

The survey was mailed between February and June of 2004. Nonrespondents were re-contacted one time only, about 1 month after the first contact, by mailing another copy of survey with a request to consider completing the survey. No phone follow-ups for completing the survey were made. This study was reviewed and approved by the University of Iowa Institutional Review Board.

Outcome Measures

The primary outcomes of the study included HRQL values, obtained through a VAS evaluation, for a "typical" health state for each of the following three conditions of oral clefting among

infants (0 to 2 years), children (3 to 12 years), adolescents (13 to 19), and adults (20 years and older): isolated cleft lip only (CLO); isolated cleft lip and palate (CLP); and a non-lethal form of syndromic clefting (SCLP). CLO and CLP were defined as unilateral or bilateral clefts of the lip and of the lip and palate, respectively, that excluded cases with recognized syndromes, cases with a chromosome abnormality, cases with one or more other major structural anomaly, or cases with cognitive delay (IQ or equivalent of less than 80). Examples of SCLP provided in the survey included the 22q– syndrome (also called DiGeorge syndrome or velocardiofacial syndrome) with cleft lip and palate. Because of potential differences in the age at which cleft surgeries occur among various teams, evaluators were instructed to assume the typical age at which clefts are repaired in their clinics.

An independent 10-cm VAS was provided for each cleft condition and age category (a total of 12 scales) to lessen any potential context measurement bias that may result from using a single scale for more than one health state. The scale was marked with 0 (death) and 100 (perfect health) on the left and right sides, respectively. Death and perfect health can be intuitively prescored at these values and thus serve as the health states that form natural bounds for the endpoints of the scale. The scale was not internally calibrated so as to avoid limiting the choice of the evaluators to any particular value. Evaluators were instructed to place the health condition (CLO, CLP, and SCLP) being evaluated for a certain age group on the scale by marking a score that they would assign to that health condition. The HRQL values were measured by the distance between 0 and the point on the scale marked by the evaluator, divided by the overall length of the scale (i.e., 10 cm). For instance, marking a scale at distance of 7 cm from the 0 point provides a HRQL value of 0.7. About 18% of evaluators wrote a direct score on the scale instead of a mark. The numeric scores were used directly in those cases. Figure 1 provides an example of the VAS scale used in this study.

Secondary outcomes in this study included the degree of confidence reported by evaluators in providing HRQL scores for each of the age groups and potential differences in scores by specialty, geographic location, and reported degree of confidence. Respondents were asked to state on a five-category Likert scale (from “strongly agree” to “strongly disagree”) the extent that they agreed with the following statement for each of the evaluated age groups: “I felt confident in completing the HRQL evaluation.”

Results

Response Rates and Evaluator Characteristics

Of the 330 selected professionals who received the survey, 153 returned the survey, yielding a crude response rate of 46%. Among those, 142 (43%) actually completed the survey. Nine of the 142 completed evaluations were considered generally unreliable because of total insensitivity of evaluators to potential effects of studied conditions on quality of life and differences by health condition and age, where all conditions were marked between 0.99 and 1 (inclusive) for all age groups. The exclusion of these evaluations had virtually no effect on the study results. Thus, the results presented here are based on 133 interpretable evaluations and an effective response rate of 40%.

Overall, the study sample was representative of all listed cleft palate and craniofacial teams on the main characteristics available in the data mentioned above. Table 1 summarizes these characteristics for all team members available for sampling, study sample members, and completers of the survey. There were no statistically significant differences between team members who provided a reliable evaluation and those who did not, including nonresponders and responders who either did not complete the evaluation or who were considered to have provided unreliable evaluations. Among the 11 professionals who returned but did not complete the survey, 9 stated that they were not confident in completing this evaluation

(including 6 who also reported being uninvolved in direct provision of care for clefting), 1 considered all individuals with clefts to be in excellent health, and another 1 gave no reason for not completing the survey.

Among the 133 interpretable evaluators, 23% were plastic surgeons, 23% were speech therapists, 11% were nurses, 11% were orthodontists, and 34% had other specialties, including dentistry, team coordination/administration, pediatrics, and other medical or surgical specialties. Those professionals were geographically located as follows (based on the four U.S. census regions): 30% were on teams located in the Midwest, 26% were in the West, 25% were in the South, and 20% were in the Northeast. Further, 30% served on teams providing cleft palate care only and 68% served on teams providing both craniofacial and cleft palate care.

HRQL Values

Table 2 reports a summary of the HRQL scores reported for each of the clefting conditions, evaluated across age groups. The means of HRQL for CLO were 0.88, 0.93, 0.93, and 0.95 for infants, children, adolescents, and adults, respectively. The means of HRQL values for CLP were 0.78, 0.82, 0.85, and 0.89 for infants, children, adolescents, and adults, respectively. The means of HRQL values for SCLP were 0.64, 0.68, 0.70, and 0.73 for infants, children, adolescents, and adults, respectively. Overall, the reported scores for all clefting conditions were clustered toward the right tail of the scale (i.e., toward high scores), with more clustering for CLO compared to CLP and SCLP. As expected, higher HRQL values were reported for CLO compared to CLP and SCLP for each age group and for CLP compared to SCLP. Within each clefting condition, higher values were reported with higher age.

Table 3 reports the degree of confidence, based on a five-category Likert scale, of completing the HRQL evaluation for each of the four age groups, as reported by the evaluators. Overall, a lower confidence level was reported for adults, where about 66% of evaluators reported confidence (including strong confidence) in completing the HRQL evaluation, compared to 79%, 84%, and 83% for infants, children, and adolescents, respectively.

A one-way analysis of variance (ANOVA) was used to test the statistical significance of differences in means of HRQL values by available evaluator characteristics, including specialty, type and geographic location of teams, and degree of confidence in conducting HRQL evaluations. Means of HRQL values by these characteristics and related statistics are not reported here but are available from the authors. Overall, there were no statistically significant differences in the means of HRQL values for the various evaluated clefting conditions by specialty in the four age groups (except for a marginal significance for adult SCLP with $p = .09$ and a 0.19-point difference between the highest for coordinator/administrator and lowest for the “other” specialty; differences were considered statistically significant at $p < .05$). Evaluators with a specialty within the category of “other surgery,” including pediatric surgery, oral-maxillofacial surgery, otolaryngology, and neurosurgery, had the lowest means for all age groups with CLO, except adolescents, and for children and adults with CLP. Evaluators with a specialty in pediatrics had the lowest means for infants with CLP and for children with SCLP. In contrast, evaluators serving as “coordinators/administrators” of their teams had the highest means for infants, children, and adults with CLP and SCLP. Evaluators with a specialty in “other” group, which included a variety of specialties, had the highest means for children, adolescents, and adults with CLO and for adolescents with CLP.

The differences in means of HRQL values by region for infants were statistically significant for CLP and SCLP (0.1-and 0.09-point difference, respectively, between highest in South and lowest in Midwest) and were marginally significant for CLO ($p = .06$, with a 0.05-point difference between highest in South and lowest in Midwest). The mean differences for adolescents with CLO were marginally significant ($p = .08$ with a 0.05-point difference

between highest in Northeast and lowest in West). Means for all age groups with CLP and SCLP and for infants with CLO were lowest for team members located in the Midwest. In contrast, the means for children, adolescents, and adults with SCLP and CLP as well as for adolescents and adults with CLO were highest for team members located in the Northeast.

Overall, the means of HRQL values for most evaluated clefting conditions and age groups were higher for members of cleft palate and craniofacial teams, compared to those of cleft-palate–only teams, with differences being statistically significant for adults with CLO and CLP (0.03- and 0.06-point differences, respectively) and marginally significant for adolescents with CLO ($p = .096$ with a 0.03-point difference) and children with CLP ($p = .06$ with a 0.04-point difference). Similarly, the means for all clefting and age groups were, overall, higher among respondents who described themselves as confident in completing the HRQL evaluation ("strongly agree" and "agree" with the confidence statement) compared to un-confident responders ("disagree" and "strongly disagree"), with differences in means (by the degree of confidence) being statistically significant only for infants with CLO (0.08-point difference).

An ordinary-least-squares (OLS) regression was estimated for HRQL values for each clefting and age group separately, using as covariates the evaluator characteristics mentioned above. Detailed OLS regression results are not reported in this paper but are available from the authors. Only results with statistical significance of $p < .05$ and selected marginally significant results ($p < .1$) are summarized. Heteroskedasticity-consistent estimation of standard errors for regression coefficients was used (White, 1980).

Overall, results similar to those of the one-way ANOVA were observed in terms of statistical significance of the effects of the characteristic variables on HRQL values. Compared to plastic surgery specalists, respondents with pediatrics or "other" specialty had about 0.16-point lower HRQL values reported for adults with SCLP. For infants, being on teams located in the Midwest was associated with a 0.06- to 0.08-point decrease in reported HRQL value for CLP compared to teams located in other regions, with a 0.05-point decrease for CLO compared to teams located in the South, and with a 0.08-point decrease for SCLP compared to teams located in the South and the West (marginally significant). For adults, members of cleft-palate–only teams reported 0.03- and 0.07-point lower HRQL values for CLO and CLP, respectively, compared to those of cleft palate and craniofacial teams. Finally, being un-confident in HRQL evaluation was associated with a 0.07-point decrease in HRQL values, compared to being confident for infants with CLO and adolescents with CLP (marginally significant).

Sensitivity Analyses

A high proportion of extreme HRQL values was reported in this sample: 46 respondents (35%) reported a perfect health HRQL score (between 0.99 and 1) for at least one clefting and age condition. Therefore, analyses were conducted to gauge the sensitivity of the study results to assumptions of overestimated HRQL and provide HRQL estimates that can also be used in sensitivity analyses of cost-effectiveness studies to validate their results for potential errors in elicitation of HRQL values. This was done by selecting respondents who gave HRQL values below certain selected ceiling HRQL values for a particular clefting/age group and then recalculating the mean of the values for this group. Three ceiling values of 0.99, 0.95, and 0.9 were selected. These values were arbitrarily chosen, yet it is hoped that this complementary analysis helps bound the "real" HRQL values and thus provide, along with the main HRQL estimates (Table 2), an estimated interval or range for these values specifically for use in cost-effectiveness studies. Of course, the results reported in Table 2 remain the best single-point estimates of the HRQL values in this study. Table 4 reports the means and standard deviations of the HRQL values for clefting and age groups under each ceiling value. More respondents reported extreme values (based on ceiling HRQL values) for adults and for higher age groups for each of the clefting conditions, but the highest proportion of extreme values was observed

for CLO. Consequently, the biggest decline in the means of HRQL values with lower ceiling values was observed for adults and for the CLO group overall.

Discussion

This work makes available HRQL values related to various states of oral clefting through age for potential use in evaluations of effectiveness and cost effectiveness of prevention and treatment interventions. The sample of evaluators, including those who responded to the survey with reliable evaluations, was of acceptable size and seemed to be, overall, representative of other potential evaluators regarding specialty, team type, and team location, providing some assurance against selection biases in reported HRQL values. The effective response rate of 40% is generally acceptable and reinforces the feasibility of research via surveying health professionals involved in care for craniofacial anomalies, including oral clefts.

The majority of evaluators reported confidence in completing the HRQL evaluation, with significantly lower confidence in evaluating HRQL for adults compared to the other age groups. The lower confidence in evaluating HRQL for adults may be a result of the fact that for most cases with oral clefting in the United States, the care provided by many members of the craniofacial teams may be concluded before adulthood. Also, the potentially lower degree of confidence reported for infants compared to children and adolescents may be a result of the complexity of defining quality of life for infants. The indicator for degree of confidence supposedly reflects overall differences in knowledge about the health condition being evaluated, as well as understanding of the HRQL concept and the evaluation task. In this sample, higher confidence levels seemed to imply higher reported HRQL values for some clefting and age groups; yet overall, differences were of a small magnitude. If the confidence measure represented instead a more optimistic view of the HRQL relating to clefting, and if nonresponse to the survey was substantially attributed to less confidence (or with a less optimistic view in this sense), the HRQL values obtained in this survey are potentially biased. It is impossible to further investigate this with the available data.

The HRQL values obtained in this survey seem to suggest a generally high HRQL associated with oral clefting and thus a low burden of clefting on individual health and quality of life. The clustering of HRQL scores toward high values may suggest a common view among the majority of evaluators that clefting imposes a limited burden on individual quality of life. However, this seems to contradict the wide clinical and research evidence and experience that emphasize the challenges introduced by this defect for all age groups, even after accounting for the fact that HRQL valuation is subjective by definition. For comparison purposes, average HRQL scores of 0.79 and 0.29 have been reported for acute otitis media and terminal-stage breast cancer, respectively, by health professionals using a VAS method (de Koning et al., 1991; Oh et al., 1996).

Another explanation for this result may be evaluators' incomplete understanding of the HRQL construct and consequently the evaluation task. Health professionals may have expressed more their perception of the direct outcomes of the care they provide for individuals with oral clefting (such as good facial appearance postsurgery or speech improvement after speech therapy) and evaluated less the overall persistent impact of clefting on quality of life, conditioned on the effectiveness and quality of provided care. The use of health professionals to evaluate HRQL has the advantage of extensive evaluator familiarity with the health condition and its impacts. This advantage may be discounted, however, with potentially biased and specifically overestimated HRQL values. This supports previous discussions in the literature about the emphasis of professionals on evaluating functional status compared to the overall status, which includes psychological, behavioral, economic, and social performance (Gold et al., 1996). In fact, the sample of evaluators, although it involved several specialties and disciplines involved

in care for clefting, was strongly underpopulated with professionals who may emphasize more the overall outcome, such as psychologists and behavioral scientists.

VAS has been criticized for its potential for bias in the measurement of HRQL values. Torrance et al. (2001) summarized these as context and end-aversion biases. *Context bias* refers to the potential effect on the evaluator's judgment of having more than one health state evaluated simultaneously. *End-aversion bias* relates to the potential tendency of evaluators to mark the scale away from reference (bounding) states (i.e., avoid both ends of the scale) and toward the middle of the scale. There is generally no strong reason to suspect different extents of aversion between the two ends of the scale, depending on the particular condition being evaluated. Providing different scales for the different evaluated health states, as was done in this study, is expected to lower the context bias. It remains possible that the specific order of presenting the health conditions in the survey (CLO-CLP-SCLP) may have introduced some ranking bias into evaluators' judgment. However, context bias cannot explain the clustering of the HRQL values toward the right tail, because this was observed for all clefting conditions and age groups. On the other hand, end-aversion bias potentially associated with VAS would lead to scores that are away from the reference states of death and perfect health and potentially to lower scores. It is thus less expected that the VAS method would have contributed to the high HRQL values observed in this study.

In general, methods that require evaluators to consider tradeoffs and uncertainty risks, such as TTO and SG, may lead to higher HRQL values than direct rating methods as a result of risk aversion of the evaluators. The use of other HRQL assessment methods to solicit preferences from this group will reveal how VAS-elicited HRQL values for oral clefting compare to those based on other methods. Further research is needed to provide a better understanding of the preferences of health professionals with respect to oral clefting—specifically, their tendency to perceive a minimal burden of this condition on health and quality of life. It is also important to directly solicit preferences of other evaluator groups, including patients with clefting (and parents of infants) and perhaps general community members, using various methods, including SG and/or TTO as well as generic instruments (e.g., Quality of Well Being Scale, Health Utilities Index) with already linked health state preferences to support multiperspective cost-effectiveness analyses.

The identity of evaluators, regardless of the evaluation method, has been an unresolved question. Different types of evaluators have been used in research, including patients, parents (or relatives), community sample members, and health professionals. Arguments for and against the appropriateness of using different types of evaluators are found in the literature, without any clear consensus (e.g., Boyd et al., 1990; Gold et al., 1996; Dolan, 1999). The optimal choice of specific evaluator groups (e.g., patients versus community members) may be dependent on the purpose of the particular cost-effectiveness analysis conducted, particularly whether it is to guide overall allocation of health care resources or is focused on treating a specific condition in a cost-effective manner (see Gold et al., 1996). A main limitation of eliciting the preferences of health professionals is that they may not be representative of those of other groups, including patients. Empirical research has usually identified differences between HRQL ratings of health professionals and other groups supporting this concern (e.g., Saigal et al., 1999). Further research is needed to identify and understand potential differences in preferences toward clefting between patients and health care providers. This would be of high clinical significance to enhance patient-provider relationships and patient outcomes for this condition.

Providing an example of VASs and scores on previously evaluated health conditions seems to influence the responses of evaluators. While this approach may improve the judgment of evaluators (lower measurement error from one side), it may introduce systematic biases in

evaluation (context bias). The net effect is likely dependent on the particular example scale and health condition being evaluated. For this study, the inclusion of an example scale in the pilot survey on acute otitis media and terminal-stage breast cancer disease with HRQL scores of 0.79 and 0.29, respectively (de Koning et al., 1991; Oh et al., 1996), led to HRQL scores that were lower by about 0.1 point (on a scale from 0 to 1) overall compared to the main study scores for the various clefting and age groups. Further investigation is needed of the effects of the use of example VASs, specifically regarding the effects of using different example scales for a specific health condition.

The objectives of this study excluded the evaluation of team characteristics, but the specialty composition of team members enrolled in ACPA seems interesting descriptively. Based on the data, a small percentage of team members were pediatricians (3.3%), and psychologists, audiologists, counselors, and social workers formed 1% or less each of all team members (Table 1). The specialty composition of team members does not seem to have been described in new studies. Strauss (1998), using similar data for 1996, described team organization and standards but did not describe the specialties of team members. Regardless of any probable inaccuracies and deficiencies in this data, including biased representation of the specialty distribution of all team members, it seems that there may be an underrepresentation of specialties, such as psychology and other social work, whose role is becoming more and more emphasized as the health burden of oral clefting is unraveled. This suggests the importance of further health services research in this area.

Conclusion

Scores obtained in this study may overestimate the real HRQL of oral clefting, given the well-documented health burden of this condition. Because of potential measurement errors (e.g., a misunderstanding of the evaluation task), possible sample selection bias (e.g., greater response to survey among more optimistic evaluators as well as differences in preferences of professionals and other groups), and our inability to effectively verify these limitations, these scores may represent the “second best” scores to use in economic (and effectiveness) evaluations related to this condition. When used in this regard, it may be useful to also perform sensitivity analyses, using the HRQL scores obtained from the sensitivity analysis reported in Table 4, to gauge the sensitivity of results to potential errors in measurement of HRQL and to different preferences regarding the impact of oral clefting on quality of life (less versus more optimistic preferences). Examples of using those scores in the future may include cost-effectiveness analyses of prevention strategies, such as vitamin supplementation, or treatment strategies, such as prenatal repair of oral clefting. It is critical to have accurate outcome measures available, as both quantitative and qualitative measures to assess changes are still needed to optimize the outcomes of those born with craniofacial anomalies. In addition, because it is increasingly recognized that clefting may be a condition with lifelong health implications (Christensen et al., 2004), ongoing performance measures, including HRQL, will also be required.

Acknowledgements

We are grateful to all health professionals who participated in the pilot and the main surveys. Development of the survey benefited from discussions with Ann Marie McCarthy, Ph.D., Sandra Daack-Hirsch, R.N., M.S.N., John Canady, M.D., and Deborah Strike, R.N., at the University of Iowa. Amy O’Reilly and Susie McConnell provided invaluable administrative support.

References

Bjornsson A, Agustsdottir S. A psychosocial study of Icelandic individuals with cleft lip or cleft lip and palate. Cleft Palate J 1987;24:152–157. [PubMed: 3472689]

Boyd NF, Sutherland HJ, Heasman KZ, Tritchler DL, Cummings BJ. Whose utilities for decision analysis? Med Decis Making 1990;10:58–67. [PubMed: 2182963]

Christensen K, Juel K, Herskind AM, Murray JC. Long-term follow-up study of survival associated with cleft lip and palate at birth. Br Med J 2004;328:1405–1408. [PubMed: 15145797]

de Koning HJ, van Ineveld BM, van Oortmarssen GJ, de Haes JC, Collette HJ, Hendriks JH, van der Maas PJ. Breast cancer screening and cost-effectiveness: policy alternatives, quality of life considerations and the possible impact of uncertain factors. Int J Cancer 1991;49:531–7. [PubMed: 1917154]

Dolan P. Whose preferences count? Med Decis Making 1999;19:482–486. [PubMed: 10520686]

Eisenberg JM. Clinical economics. A guide to the economic analysis of clinical practices. JAMA 1989;262:2879–2886. [PubMed: 2509745]

Froberg DG, Kane RL. Methodology for measuring health-state preferences–II: Scaling methods. J Clin Epidemiol 1989;42:459–471. [PubMed: 2732774]

Gold, MR.; Patrick, DL.; Torrance, GW.; Fryback, DG.; Hadorn, DC.; Kamlet, MS.; Daniels, N.; Weinstein, MC. Identifying and valuing outcomes. In: Gold, MR.; Russell, LB.; Siegel, JE.; Weinstein, MC., editors. Cost-Effectiveness in Health and Medicine. New York: Oxford University Press; 1996. p. 82-134.

Heller A, Tidmarsh W, Pless IB. The psychosocial functioning of young adults born with cleft lip or palate. A follow-up study. Clin Pediatr (Phila) 1981;20:459–465. [PubMed: 7237883]

Jelliffe-Pawlowski LL, Shaw GM, Nelson V, Harris JA. Risk of mental retardation among children born with birth defects. Arch Pediatr Adolesc Med 2003;157:545–550. [PubMed: 12796234]

Jocelyn LJ, Penko MA, Rode HL. Cognition, communication, and hearing in young children with cleft lip and palate and in control children: a longitudinal study. Pediatrics 1996;97:529–534. [PubMed: 8632941]

Jones MC. Etiology of facial clefts. Prospective evaluation of 428 patients. Cleft Palate J 1988;25:16–20. [PubMed: 3422594]

Kapp-Simon KA, Simon DJ, Kristovich S. Self-perception, social skills, adjustment, and inhibition in young adolescents with craniofacial anomalies. Cleft Palate Craniofac J 1992;29:352–356. [PubMed: 1643066]

Kapp-Simon K. Self-concept of primary-school-age children with cleft lip, cleft palate, or both. Cleft Palate J 1986;23:24–27. [PubMed: 3455899]

Krabbe PF, Essink-Bot ML, Bonsel GJ. The comparability and reliability of five health-state valuation methods. Soc Sci Med 1997;45:1641–1652. [PubMed: 9428084]

Marcusson A, Akerlind I, Paulin G. Quality of life in adults with repaired complete cleft lip and palate. Cleft Palate Craniofac J 2001;38(4):379–385. [PubMed: 11420018]

Mossey, PA.; Little, J. Epidemiology of oral clefts: an international perspective. In: Wyszynski, DF., editor. Cleft Lip and Palate: From Origin to Treatment. 2002. New York: Oxford University Press; p. 127-158.

Nackashi, JA.; Dedlow Rosellen, E.; Dixon-Wood, V. Health care for children with cleft lip and palate: comprehensive services and infant feeding. In: Wyszynski, DF., editor. Cleft Lip and Palate from Origin to Treatment. New York: Oxford University Press; 2002. p. 127-158.

Nopoulos P, Berg S, VanDemark D, Richman L, Canady J, Andreasen NC. Cognitive dysfunction in adult males with non-syndromic clefts of the lip and/or palate. Neuropsychologia 2002;40:2178–2184. [PubMed: 12208013]

Oh PI, Maerov P, Pritchard D. A cost-utility analysis of second line antibiotics in the treatment of acute otitis media in children. Clin Ther 1996;18:160–82. [PubMed: 8851461]

Ramstad T, Ottem E, Shaw WC. Psychosocial adjustment in Norwegian adults who had undergone standardised treatment of complete cleft lip and palate. I. Education, employment and marriage. Scand J Plast Reconstr Surg Hand Surg 1995a;29:251–257. [PubMed: 8539569]

Ramstad T, Ottem E, Shaw WC. Psychosocial adjustment in Norwegian adults who had undergone standardised treatment of complete cleft lip and palate. II. Self-reported problems and concerns with appearance. Scand J Plast Reconstr Surg Hand Surg 1995b;29:329–336. [PubMed: 8771260]

Read JL, Quinn RJ, Berwick DM, Fineberg HV, Weinstein MC. Preferences for health outcomes. Comparison of assessment methods. Med Decis Making 1984;4:315–329. [PubMed: 6335216]

Saigal S, Stoskopf BL, Feeny D, Furlong W, Burrows E, Rosenbaum PL, Hoult L. Differences in preferences for neonatal outcomes among health care professionals, parents, and adolescents. JAMA 1999;281:1991–1997. [PubMed: 10359387]

Strauss RP. Cleft palate and craniofacial teams in the United States and Canada: a national survey of team organization and standards of care. The American Cleft Palate-Craniofacial Association (ACPA) Team Standards Committee. Cleft Palate Craniofac J 1998;35:473–480. [PubMed: 9832217]

Thomas PT, Turner SR, Rumsey N, Dowell T, Sandy JR. Satisfaction with facial appearance among subjects affected by a cleft. Cleft Palate Craniofac J 1997;34:226–231. [PubMed: 9167073]

Torrance GW. Measurement of health state utilities for economic appraisal: a review. J Health Econ 1986;5:1–30. [PubMed: 10311607]

Torrance GW, Feeny D. Utilities and quality-adjusted life years. Int J Technol Assess Health Care 1989;5:559–575. [PubMed: 2634630]

Torrance GW, Feeny D, Furlong W. Visual analog scales: do they have a role in the measurement of preferences for health states? Med Decis Making 2001;21:329–334. [PubMed: 11475389]

White H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 1980;48:817–838.

FIGURE 1.
Visual analogue scale used to obtain HRQL values for oral clefting by age.

TABLE 1

Characteristics of Study Population, Sample, and Responding Group of Team Members

Characteristic	*No. (%)*		
	Population (N = 1334)	***Survey Sample (N = 330)***	***Survey Completers (N = 133)***
Specialty*			
Nursing	123 (9.3)	43 (13.1)	15 (11.4)
Speech therapy	185 (14.0)	58 (17.7)	30 (22.7)
Dentistry†	62 (4.7)	18 (5.5)	7 (5.3)
Orthodontics	153 (11.5)	37 (11.3)	15 (11.4)
Plastic surgery	377 (28.4)	85 (26.0)	31 (23.5)
Other surgery‡	211 (15.9)	30 (9.2)	14 (10.6)
Pediatrics	44 (3.3)	5 (1.5)	3 (2.3)
Coordination/administration	86 (6.5)	27 (8.2)	10 (7.6)
Other§	85 (6.4)	25 (7.6)	7 (5.3)
Team region‖			
Northeast	237 (17.8)	55 (16.7)	26 (19.6)
Midwest	371 (27.8)	84 (25.5)	40 (30.1)
South	405 (30.4)	86 (26.1)	33 (24.8)
West	321 (24.1)	105 (31.8)	34 (25.6)
Type of team			
Cleft palate only	466 (34.9)	107 (32.4)	40 (30.1)
Both cleft palate and craniofacial	798 (59.8)	215 (65.2)	90 (67.7)
Other¶	70 (5.3)	8 (2.43)	3 (2.3)

*Fewer than 0.7% of study population of team members, survey recipients, and survey completers who did not have the specialty characteristic available.

†Includes general dentistry, pediatric dentistry, and prosthodontics/prosthetics.

‡Includes pediatric surgery, oral-maxillofacial surgery, otolaryngology, and neurosurgery.

§Includes social work, counseling, dietetics, nutrition, dental hygiene, occupational therapy, research, psychology, anatomy, audiology, and genetics.

‖Based on the U.S. census classification of states into the four geographic regions.

¶Includes craniofacial teams only and teams that were not specified/reported into one of the above categories.

TABLE 2

Means, Standard Deviations, Medians, 25th Percentiles, and Minimum/Maximum Values of Obtained HRQL Scores for CLO, CLP, and SCLP by Age

Clefting/Age Group	*Sample Size*[*]	*Mean*	*Standard Deviation*	*Median*	*25th Percentile*	*Minimum*	*Maximum*
CLO[†]							
Infants	131	0.88	0.10	0.90	0.85	0.48	1.00
Children	132	0.93	0.07	0.95	0.90	0.54	1.00
Adolescents	133	0.93	0.08	0.95	0.90	0.59	1.00
Adults	125	0.95	0.07	0.97	0.94	0.60	1.00
CLP							
Infants	131	0.78	0.13	0.80	0.72	0.23	1.00
Children	132	0.82	0.13	0.85	0.74	0.30	1.00
Adolescents	133	0.85	0.12	0.87	0.77	0.50	1.00
Adults	125	0.89	0.10	0.90	0.85	0.49	1.00
SCLP							
Infants	129	0.64	0.16	0.67	0.56	0.23	0.93
Children	128	0.68	0.16	0.69	0.61	0.22	0.98
Adolescents	130	0.70	0.16	0.71	0.60	0.25	0.99
Adults	122	0.73	0.17	0.76	0.63	0.23	0.99

[*] Some responders did not complete the evaluation for all age groups and/or clefting conditions, resulting in sample sizes of less than 133 for some clefting/age groups.

[†] CLO = isolated cleft lip only, CLP = isolated cleft lip and palate, SCLP = nonlethal form of syndromic clefting.

Cleft Palate Craniofac J. Author manuscript; available in PMC 2006 November 21.

TABLE 3

Confidence Level for Completing HRQL Evaluation by Age

	No. (%)			
Confidence Level	***Infants***	***Children***	***Adolescents***	***Adults***
Strongly confident	57 (44.5)	58 (45.3)	51 (39.8)	31 (24.8)
Confident	44 (34.4)	50 (39.1)	55 (43.0)	52 (41.6)
Neutral	14 (10.9)	10 (7.8)	13 (10.2)	26 (20.8)
Unconfident	8 (6.3)	7 (5.5)	7 (5.5)	10 (8.0)
Strongly unconfident	5 (3.9)	3 (2.3)	2 (1.6)	6 (4.8)

TABLE 4

Means and Standard Deviations (SD) of HRQL Scores for CLO, CLP, and SCLP by Age for Respondents with HRQL Values Below 0.9, 0.95, and 0.99

	HRQL < 0.9		*HRQL < 0.95*		*HRQL < 0.99*	
Clefting/Age Group	***N****	***Mean (SD)***	***N****	***Mean (SD)***	***N****	***Mean (SD)***
CLO†						
Infants	56	0.80 (0.09)	92	0.85 (0.09)	121	0.87 (0.09)
Children	26	0.82 (0.08)	64	0.88 (0.07)	106	0.91 (0.07)
Adolescents	24	0.80 (0.08)	61	0.87 (0.08)	102	0.91 (0.08)
Adults	12	0.79 (0.09)	34	0.87 (0.08)	87	0.93 (0.07)
CLP						
Infants	108	0.75 (0.12)	123	0.77 (0.12)	127	0.7 (0.13)
Children	86	0.75 (0.11)	113	0.79 (0.12)	128	0.81 (0.13)
Adolescents	79	0.77 (0.10)	99	0.80 (0.11)	120	0.83 (0.12)
Adults	55	0.81 (0.09)	81	0.84 (0.09)	108	0.87 (0.10)
SCLP						
Infants	124	0.63 (0.16)	129	0.64 (0.16)	129	0.64 (0.16)
Children	118	0.66 (0.15)	124	0.67 (0.15)	128	0.68 (0.16)
Adolescents	110	0.66 (0.14)	123	0.68 (0.15)	129	0.70 (0.16)
Adults	101	0.70 (0.15)	111	0.71 (0.16)	120	0.73 (0.16)

*Relates to the total number of respondents who reported HRQL values below the selected cutoff value for a particular clefting/age group.

†CLO = isolated cleft lip only, CLP = isolated cleft lip and palate, SCLP = nonlethal form of syndromic clefting.

Cleft Palate Craniofac J. Author manuscript; available in PMC 2006 November 21.