Orthodontics & Craniofacial Research

ORIGINAL ARTICLE

A. J. Akram
A. J. Ireland
K. C. Postlethwaite
J. R. Sandy
A. S. Jerreat

Assessment of a condition-specific quality-of-life measure for patients with developmentally absent teeth: validity and reliability testing

Authors' affiliations:
A. J. Akram, A. S. Jerreat, Department of Orthodontics, Royal Devon and Exeter NHS Foundation Trust, Exeter, UK
A. J. Ireland, J. R. Sandy, School of Oral and Dental Science, Bristol, UK
K. C. Postlethwaite, Department of Education, Exeter University, Exeter, UK

Correspondence to:
Ansa Akram
Orthodontic Department
Royal Devon and Exeter NHS Foundation Trust
Barrack road
Exeter
Devon, UK
E-mail: Ansa_akram@yahoo.com

Akram A. J., Ireland A. J., Postlethwaite K. C., Sandy J. R., Jerreat A. S. Assessment of a condition-specific quality-of-life measure for patients with developmentally absent teeth: validity and reliability testing *Orthod Craniofac Res* 2013; 16: 193–201. © 2013 John Wiley & Sons A/S. Published by Blackwell Publishing Ltd

Structured Abstract

Objectives – This article describes the process of validity and reliability testing of a condition-specific quality-of-life measure for patients with hypodontia presenting for orthodontic treatment. The development of the instrument is described in a previous article.

Setting and Sample population – Royal Devon and Exeter NHS Foundation Trust & Musgrove Park Hospital, Taunton.

Materials and Methods – The child perception questionnaire was used as a standard against which to test criterion validity. The Bland and Altman method was used to check agreement between the two questionnaires. Construct validity was tested using principal component analysis on the four sections of the questionnaire. Test–retest reliability was tested using intraclass correlation coefficient and Bland and Altman method. Cronbach's alpha was used to test internal consistency reliability.

Results – Overall the questionnaire showed good reliability, criterion and construct validity. This together with previous evidence of good face and content validity suggests that the instrument may prove useful in clinical practice and further research.

Conclusions – This study has demonstrated that the newly developed condition-specific quality-of-life questionnaire is both valid and reliable for use in young patients with hypodontia.

Key words: hypodontia; quality of life; questionnaire; reliability; validity

Date:
Accepted 30 December 2012

DOI: 10.1111/ocr.12018

Introduction

In 1992, the UK Department of Health proposed that health status and health-related quality of life (HRQoL) should be incorporated into outcome assessment along with survival rates, symptoms and complications (1). The importance of quality of life in the field of dentistry has been acknowledged only relatively recently (2).

Traditionally, the outcome of orthodontic treatment is assessed using measures such as the PAR (peer assessment rating) score or changes in cephalometric measurements. These measures reflect only the professional perception of outcome and take no account of the disability experienced by the patient or consumer of care. Similarly, the Index of Orthodontic Treatment Need (IOTN) is primarily a clinician-based measure of normative rather than patient perceived need (3). At present, there is no single standard condition-specific OHRQoL measure for use in orthodontics. Most of the available measures designed to assess oral health-related quality of life (OHRQoL) are unsuitable for use in orthodontic patients (4).

In contrast to their peers, patients with hypodontia undergo complex and lengthy multidisciplinary treatment commencing at a very young age. As a result, the issues of importance for them can be very different from patients undergoing routine orthodontic treatment. This means that existing OHRQoL instruments for assessing quality of life are of limited use in this group of patients. It was for this reason a condition-specific quality-of-life questionnaire was developed for patients with hypodontia, the development of which has been described in a previous article (5). The aim of this present study was to assess the internal consistency reliability, test–retest reliability, criterion and construct validity of this newly developed instrument. Key indicators of the quality of a measuring instrument are the reliability and validity of the measure.

According to classical test theory, any score obtained by a measuring instrument is composed of both the 'true' score which is unknown and the 'error' in the measurement process. The true score is essentially the score that a person would have received if the measurement was perfectly accurate. The process of validating an instrument is in large part focused on reducing error in the measurement process (6). Reliability estimates are used to evaluate the stability of measures administered at different times to the same individual (test–retest reliability) or equivalence of sets of items from the same test (internal consistency).

The test–retest reliability of the method evaluates the extent to which the observer obtains the same results in identical circumstances. This is undertaken by repeated administration of the questionnaire. The responses are then compared on both occasions (7). An appropriate time interval needs to be selected: too long and things may have changed; too short and patients may remember their first response and put it down, rather than answering the question *de novo*.

Internal consistency reliability assesses the consistency of results across items within a test. All of the items should be tapping different aspects of the same attribute and not different parts of different traits. Current thinking in test development holds that there should be a moderate correlation between the items in a scale (8). If the items are chosen without regard for homogeneity, then the resulting scale could possibly end up tapping a number of traits.

Criterion validity refers to the correlations of the questionnaire with another criterion measure, which is accepted as the gold standard. It may not be possible to test criterion validity if no gold standard is available (9). The most commonly used design for testing criterion validity is to administer the new scale and the standard at the same time. The results can be analyzed by using a measure of correlation, this is termed concurrent criterion validity.

Construct validity is the extent to which the instrument tests the hypothesis or theory it is measuring. Construct validity is particularly important when we move away from the realm of physical attributes into more psychological ones like anxiety, intelligence or pain. We cannot directly observe these variables (8). The proposed underlying variables are referred to as

hypothetical constructs. A construct can be thought of as a 'mini-theory' to explain the relationships between various behaviors or attitudes. For instance, patients with absent anterior teeth would be expected to score higher on the impact on appearance score than patients who are, for example, missing second premolars only.

Materials and methods

Study sample

Ethical approval to test the validity and reliability of the newly developed questionnaire was granted by the Devon and Torbay Research and Ethics Committee (REC 09/H0202/24). A total of 55 consecutive patients with hypodontia aged between 11 and 18 years were recruited into the study from two orthodontic units in the south-west of England. This included patients at different stages of treatment and with varying degrees of severity of hypodontia. Patients with cleft lip and palate and patients with hypodontia as part of a syndrome were excluded.

Information sheets about the questionnaire were provided to the patients and parents/guardians, and written informed consent obtained before the questionnaire was completed. Patients were asked to complete the questionnaire themselves in the waiting room during one of their routine appliance adjustment appointments. The questionnaire was divided into four sections, namely treatment, activities, appearance and the reaction of other people. The aim of the study was to test the criterion validity, construct validity and reliability on a large sample of patients.

Determination of criterion validity

Criterion validity refers to the correlation of the questionnaire with another criterion measure, which is accepted as the gold standard. No such gold standard measure is available for patients with hypodontia. Therefore, the child perception questionnaire (CPQ) (10) was used as a substitute against which to test criterion validity. Although the CPQ is a generic instrument and has significant limitations, which have been highlighted in recent studies ((11)-(12)), it was designed to be used in a similar age group of patients and has been used in previous studies on OHRQoL. Not all of the sections and questions within the CPQ were relevant to the current research question. Therefore, questions in the CPQ and the hypodontia questionnaire, which would be expected to have a high correlation, were selected independently by two researchers. The selected questions were then compared and agreement was reached. The participants were asked to fill in the CPQ and hypodontia questionnaire at the same visit.

Responses from the two questionnaires were entered into SPSS© software v16.0. Correlations were compared for the responses using Spearman's rank correlation coefficient (13). The Spearman's coefficient is a nonparametric measure of statistical dependence between two variables and can be used when the data is non-normally distributed.

To investigate the agreement between the two measures, the Bland and Altman method was used. This plots the mean of the two measurements on the *X*-axis and the difference between the two measurements on the *Y*-axis (14). Limits of agreement are set as mean $\pm$ 2 standard deviations (SD).

Determination of Construct validity

Principal components analysis (PCA) was used to investigate the underlying dimensionality of the data and to determine whether the instrument was made of subscales. Components are made up of groups of correlated questions. In this method, to decide which components were statistically important, 'eigenvalues' were calculated for each component. A graph called a scree plot of each eigenvalue (*Y*-axis) against the component with which it is associated (*X*-axis) was plotted using a facility in SPSS©. It is recommended to retain all components with eigenvalues greater than 1(15). Principal component analysis also assigns rotated matrix scores to each question in the component. The higher these scores, the more closely related the question is to the component.

Determination of reliability

To assess the test–retest reliability, the hypodontia questionnaire was administered on two occasions a minimum of 4 weeks apart (T1 and T2). The researcher ensured that no significant clinical change had taken place between T1 and T2, such as removal of the fixed appliances or replacement of a missing tooth. The responses from T1 and T2 were entered into SPSS©, and the intraclass correlation coefficient (ICC) was calculated. A Bland and Altman test was also carried out between T1 and T2 to investigate the agreement between the responses.

Cronbach's alpha (α) was carried out on the questionnaire as a whole and on the different sections of the questionnaire to test internal consistency. The value of Cronbach's alpha is dependent on the number of items, that is, the higher the number of items the higher the value of alpha, and this is sometimes quoted as a disadvantage of using this approach (16). Keeping this in mind, Cronbach's alpha was not calculated for the treatment section of the questionnaire as there were only three items in this section.

Results

Criterion validity

Of the 55 patients recruited, 46 patients agreed to take part in this section of the study, 14 were male and 32 female. The age range was still 11–18 years. The criterion validity results for Spearman's correlation (þ (rho)) are shown in Table 1. The results show that there were moderate to high correlations between the seven questions from the hypodontia questionnaire and the seven questions that were selected from the CPQ (p = 0.4–0.7, p < .01).

The Bland and Altman plots for the seven items showed that overall agreement between the responses was good, at 80% or more. The extent of agreement for the scores was similar across the range of scores, that is, participants with low or high mean scores agreed similarly and there were no trends.

Figure 1 is an example of a Bland and Altman plot for item 3. This shows that 94% of responses were within the normally used limits of agreement, that is, 2 SD from the mean.

Construct validity

Principal component analysis was carried out on the four sections of the questionnaire to assess if there were sub divisions within the sections. Figure 2 shows the scree plot for the treatment section of the questionnaire. There were three questions in the treatment section that could be included in PCA. As can be seen from the graph in Fig. 2, only one component had an eigenvalue above 1.00, and therefore, there was only one significant component in this section. Table 2 shows

Table 1. **Results of Spearman's rank correlation coefficient for criterion validity**

Hypodontia questionnaire	CPQ	Correlation
I feel embarrassed about the way my teeth look?	How satisfied are you with the appearance of your teeth?	Moderate
Food gets stuck in the gaps between my teeth?	Food stuck in between your teeth?	Moderate
Having missing teeth affects my speech?	Difficult to say any words?	High
I feel embarrassed about meeting people for the first time because of the way my teeth look?	Felt shy or embarrassed?	Moderate
Most of my friends have better looking teeth than I do?	Worried that you are not as good looking as others?	Moderate
I don't laugh out loud with friends because of the way my teeth look?	Avoid smiling or laughing when around other children?	Moderate
People take the mickey out of me because of the way my teeth look?	Other children teased you or called you names?	Moderate

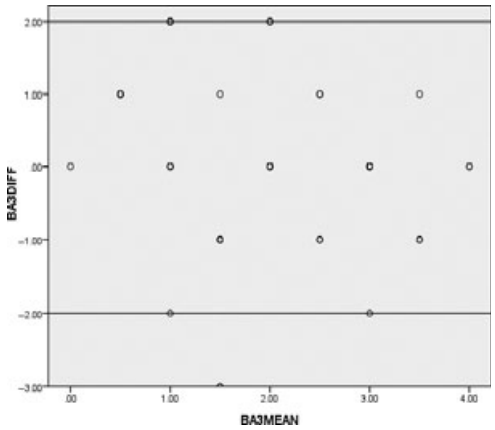

Fig. 1. Bland and Altman plot for item 3.

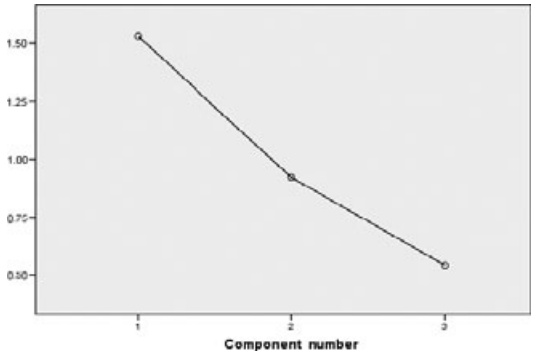

Fig. 2. Scree plot for treatment section.

Table 2. **Rotated matrix scores for questions in component 1 of treatment section**

Questions	Component 1
Treatment complexity	0.839
Length of treatment	0.738
Worries about end result	0.532

that all three questions in the treatment section were in component 1. This means that there were no subdivisions within this section and that the questions were correctly grouped together.

The scree plot for the activities section (Fig. 3) shows that there were two components in this section with eigenvalues above 1.00. Table 3 shows that six of the seven questions in this section were in component one. The second component had an eigenvalue of just above 1.00, and there was only one question in this component that was not in component one. It is usual to discard components which are associated with only one or two variables (Klein, 1994) (17); therefore, this subsection also appears to be unidimensional.

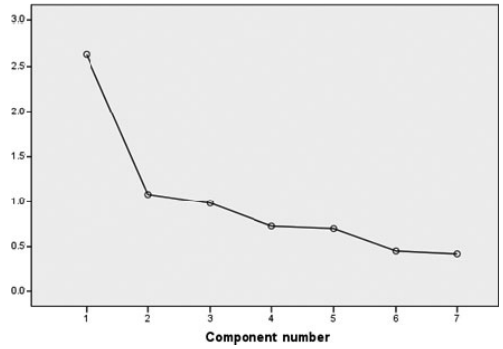

Fig. 3. Scree plot for activities section.

Table 3. **Rotated matrix scores for components 1 and 2 in the activities section**

	Component	
Questions	1	2
Tooth brushing	0.776	
Food stuck in gaps	0.701	
Speaking out aloud	0.674	
Contact sports	0.605	
Effects on speech	0.601	0.488
Don't eat in public	0.596	−0.455
Difficulty chewing foods?		0.630

The scree plot for the appearance section of the questionnaire also showed that there were two components with eigenvalues above 1.00 with six of the seven questions in this section being placed in component one.

In the other people's reaction section of the questionnaire, the plot (Fig. 4) showed that there were three components in this section with eigenvalues above 1.00. Table 4 shows that five of the seven questions in this section were in component one. It is evident from the table that questions tapping into the family aspects were grouped as a separate component from those tapping into the friends and peers aspects.

Reliability testing results

Ten patients were asked to participate in this part of the study, two were male and eight were female. Table 5 shows the ICC's for the test–retest reliability. The Bland and Altman plot (Fig. 5) shows that 81% of responses for T1 and T2 were within the limits of agreement set at 2

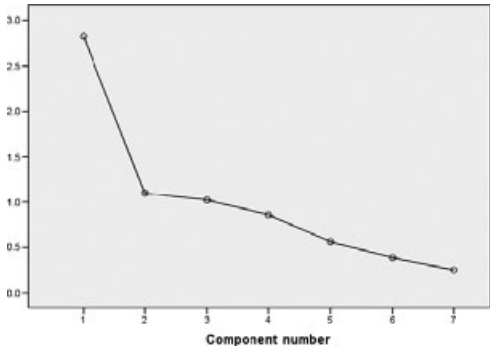

Fig. 4. Scree plot for reaction of other peoples' section.

Table 4. **Rotated matrix scores for questions in components 1,2 and 3 of the other peoples' reaction section**

	Component	
Questions	1	2
Gaps bother me	0.891	
Embarrassed about teeth	0.860	
Photographs	0.757	
Friends have better looking teeth	0.683	
Out of proportion	0.663	
Laughing out aloud with friends	0.651	−0.493
Teeth are smaller than friends		0.860

Table 5. **Results of the test–retest reliability of the questionnaire using an intraclass correlation coefficient**

Questionnaire section	Intraclass correlation coefficient
Treatment	0.91
Activities	0.88
Appearance	0.97
Other people	0.77

standard deviations from the mean. It also shows that people with average scores, that is, 0 or 4 tend to vary very little, but people with middle scores vary more.

Cronbach's alpha (α) was carried out to test the internal consistency reliability of the questions that were grouped together in the components with eigenvalues above 1.00, in each of the four sections of the questionnaire. Cronbach's alpha was carried out on component 1 in the activities, appearance and other people's reaction sections. Other components in these sections were not

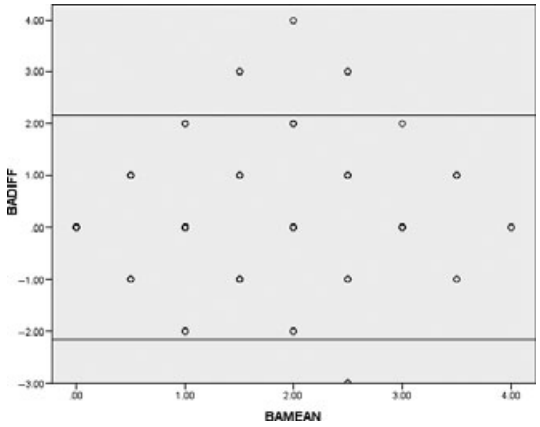

Fig. 5. Bland and Altman plot for test–retest reliability.

Table 6. **Cronbach's alpha for the overall questionnaire and three sections of the questionnaire**

Questionnaire	Cronbach's alpha
Overall	0.89
Activities	0.74
Appearance	0.80
Other people	0.79

considered as the majority of items in these sections were in component one. Cronbach's alpha for the overall questionnaire (Table 6) was high which shows that the individual items in the questionnaire are measuring the same trait and are highly correlated with each other. In addition, Table 6 shows that the Cronbach's alpha was acceptable for all sections of the questionnaire.

This condition-specific quality-of-life instrument is available through the database PROQOLID (www.proqolid.org).

Discussion

This condition-specific quality-of-life questionnaire for patients with hypodontia had previously been created using a rigorous process of development and testing in which it was shown to have good face and content validity (6). However, before it could be recommended for use in patients with hypodontia, it was necessary to assess its validity and reliability as outlined in the present article. The study sample that was chosen for this study excluded patients with hypodontia as part of a syndrome as it was felt that in

this group of patients, it would be difficult to differentiate between issues related to the missing teeth and those related to the syndrome.

Criterion Validity

Of the 55 patients recruited to the study, 46 agreed to complete both the CPQ and the hypodontia questionnaire. Nine declined to complete the CPQ due to time constraints. It was felt that 46 participants were an acceptable number for criterion validity testing.

Studies aimed at developing new health measurement scales often state that criterion validity tests cannot be carried out due to lack of a gold standard method, hence the need to develop a new one. Although this was true for the current study, it was decided to select items from the child perception questionnaire, which could be directly compared to items in new condition-specific hypodontia questionnaire. This was to ensure a more in depth statistical analysis. Seven items in each questionnaire were recognized, and the results (Table 1) showed a positive moderate to high correlation between all the seven items selected in both questionnaires. This suggests that scores from the new condition-specific hypodontia questionnaire correspond well with scores from the child perception questionnaire on similarly worded items. Hence for the selected items, there is an acceptable level of correlation between the new measure and the previously used standard measure.

The Bland and Altman plots in each case showed no obvious pattern, that is, there was a random distribution. The percentage agreement between the two questionnaires for the seven items was good with a low percentage of outliers outside the limits of agreement. There was no trend to the agreement which means that agreement was consistent between the two questionnaires and was not affected by whether the score on the questions was high or low.

Construct Validity

Principal component analysis (PCA) was used to test construct validity. Principal component analysis is concerned with establishing which linear components exist within the data and how particular items relate to that component. It was carried out on all four sections of the questionnaire. The results for the treatment section (Fig. 2 and Table 2) showed that there was one component and that all three of the questions in this section, which were selected for the analysis, were related to this component. Hence, this section was unidimensional. The higher the rotated component matrix scores for the questions in the component, the stronger the relation is to the component.

In the activities section of the questionnaire, PCA revealed two components (Fig. 3 and Table 3). Six of the seven questions selected for the PCA were in component one. Component two had an eigenvalue just above 1.00, and therefore, the relative importance of this component was low. Also, one question appeared in component two and did not appear in component one. This question asks about difficulty in chewing foods. It was felt that in answering this particular question, patients found it difficult to distinguish between difficulties in chewing foods due to fixed appliances, as opposed to difficulties in chewing due to missing teeth. This question will be eliminated from the questionnaire as it is not sufficiently discriminatory.

Interestingly, PCA of the appearance section also revealed two components with eigenvalues above 1.00. Six of the seven questions in this section were related to component one. Only one question appeared in component two that did not relate to component one. This question asked about whether participants felt their teeth were smaller in size compared to their friend's teeth. Although this was mentioned by a few participants in the focus group stage in the initial study (6), it will eliminated from the final questionnaire as it is not sufficiently discriminatory. One of the reasons for this may be that some participants may find it difficult to make this judgment.

The other people's reaction section of the questionnaire revealed three components (Fig. 4 and Table 7). Five of the seven questions in this section were related to component one. Interestingly, the questions in component two and three

Table 7. **Rotated matrix scores for questions in component 1 and 2 of the appearance section**

	Component	
Questions	1	2
Gaps bother me	0.891	
Embarrassed about teeth	0.860	
Photographs	0.757	
Friends have better looking teeth	0.683	
Out of proportion	0.663	
Laughing out aloud with friends	0.651	−0.493
Teeth are smaller than friends		0.860

were those exploring relationships with family members as opposed to the questions exploring relationship with friends and peers that were part of component one. This suggests that the questionnaire could be amended by grouping the questions exploring relationships with family members, as a separate section in the questionnaire, and perhaps developing some additional questions to map this aspect of participants' views more fully.

These results offer support for the construct validity of this condition-specific instrument. The proposed underlying hypothetical constructs were largely supported by the data which showed that most of the questions in the individual sections were grouped correctly and should be retained as such.

Test–retest Reliability

The test–retest reliability of the method evaluates the extent to which the observer obtains the same results in identical circumstances. It was confirmed that no significant clinical change had taken place between T1 and T2 for the patients involved in this test, for example, removal or bonding of fixed appliances or artificial tooth replacement, as this could affect the responses.

Two separate statistical analyses were performed, an ICC and the Bland and Altman method. The ICC results were all within the strong to very strong correlation range, with values above 0.70, and thus gave no cause for concern (Table 5).

The Bland and Altman plot showed that there was no pattern to the points, that is, there was a random distribution. 81% of responses were within the limits of agreement, which were set as 2 SD from the mean (Fig. 5). Although the agreement is not as high as 95%, it is still within the good agreement range. Also, the trend in the distribution shows that patients that scored very high or very low were more consistent in their scores at T1 and T2 compared to patients that were in the middle category. This suggests that views for patients with very low or very high impact on quality of life remain relatively stable throughout treatment and are more reliable than patients with moderate impact of hypodontia.

The results therefore demonstrate that the questionnaire displays good test–retest reliability. Other studies such as Cunningham et al. (2000) (18)also used ICC to assess test–retest reliability, of a condition-specific quality-of-life measure for patients with dentofacial deformity, which the first part of their study had developed. Alternatively, McNair et al. (2009) (19)used kappa coefficients to assess test–retest reliability of a patient-based questionnaire to assess satisfaction with the process of orthodontic treatment.

Internal consistency reliability

Cronbach's alpha was used to test internal consistency reliability. Cronbach's alpha values for both the overall questionnaire and the sections of the questionnaire were high and greater than the generally accepted value of 0.7 (Kline, 1999) (12). This suggests that all the questions in the different sections of the questionnaire are consistent with each other.

Cronbach (1951) (20)suggested that if a questionnaire has subsections or subscales, α should be applied separately to the subsections as well. However, the value of α is dependent on the number of items in the subscale. The results of Cronbach's alpha are less reliable with a small number of items. Therefore, alpha was not calculated separately for the treatment section of the questionnaire as there were only three items in this section.

The value of Cronbach's alpha for the other three sections of the questionnaire was above 0.7

(Table 6). This would also suggest that the individual questions in each of these sections are consistent with each other and are correctly grouped together. Cronbach's alpha has been used in previous studies to assess internal consistency reliability of oral health questionnaires. Jokovic et al. (2002) used Cronbach's alpha to test internal consistency reliability of the child perception questionnaire (6). Slade and Spencer (1994) (21)also used Cronbach's alpha to assess reliability of the oral health impact profile questionnaire.

Conclusion

This study and the previous article (6) have demonstrated that the newly developed condition-specific quality-of-life questionnaire for patients with hypodontia is both a valid and a reliable tool. The intention is to eliminate the two questions highlighted in the construct validity section and then use the questionnaire in a large trial to assess the quality of life in patients with hypodontia. It can also be considered for use to influence issues such as funding for patients with hypodontia in the UK.

Clinical relevance

It is widely accepted that the developmental absence of teeth can impact significantly on a person's quality of life. However, there is currently a lack of robust scientific evidence to support this view and also how orthodontic treatment might improve the quality of life. It is for this reason a health-related condition-specific quality-of-life questionnaire was developed for patients with hypodontia.

Acknowledgments: We are grateful to the patients who took part in this study. We would also like to thank the orthodontic staff in the two departments for all their help and support.

References

1. Department of health. *An Oral Health Strategy for England*. London: HMSO; 1994.
2. Cunningham SJ, Hunt NP. Quality of life and its importance in orthodontics. *J Orthod* 2001;28:152–158.
3. Kok VY, Mageso P, Harradine NWT, Sprod AJ. Comparing a quality of life measure and the aesthetic component of the index of orthodontic treatment need and concern. *J Orthod* 2004;31:312–318.
4. O'Brien K, Kay L, Fox D, Mandall N. Assessing oral health outcomes for orthodontics- Measuring oral health status and quality of life. *Community Dent Health* 2004;15:22–26.
5. Akram A, Jerreat AS, Woodford J, Sandy JR, Ireland AJ. Development of a condition specific instrument to assess quality of life measure in hypodontia patients. *Orthod Craniofac Res* 2011;14:160–7.
6. Crocker L, Algina J. *Introduction to Classical and Modern Test Theory*. Orlando,FL: Harcourt Brace Jovanovich; 1986.pp. 1–52.
7. Jensen MP. Questionnaire validation: a brief guide for readers of research literature. *Clin J Pain* 2003;19:345–42.
8. Saw SM, Ng TP. The design and assessment of questionnaires in clinical research. *Sing Med J* 2001;42:131–135.
9. Carmines EG, Zeller RA. *Reliability and Validity Assessment*. London: SAGE publications; 1979.
10. Jokovic A, Locker D, Guyatt G. Short forms of the child perception questionnaires for 11-14 year old children: development and initial evaluation. *Health Qual Life Outcome* 2006;4:4.
11. Marshman Z, Gibson BJ, Benson PE. Is the short-form of child perception questionnaire meaningful and relevant to children with malocclusion in the UK? *J Orthod* 2010;37:29–36.
12. O'Brien C., Benson.P.E. , Marshman Z. . Evaluation of a quality of life measure for children with malocclusion. *J Orthod* 2007;34:184–193.
13. Spearman C. The proof and measurement of association between two things. *Int J Epidemiol* 1904;15:72–101.
14. Bland J.M, Altman D.G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–310.
15. Kaiser H.F. An index of factorial simplicity. *Psychometrika* 1974;35:401–415.
16. Cortina J.M. What is coefficient alpha? An examination of theory and applications. *J Appl Psychol* 1993;78:98–104.
17. Kline P. *An Easy Guide to Factor Analysis*. London: Routledge; 1994.
18. Cunningham S., Garratt A.M., Hunt N.P. Development of a condition specific quality of life measure for patients with dentofacial deformity: 1. Reliability of the instrument.. *Community Dent Oral Epidemiol* 2004;28:195–201.
19. Laing E, Cunningham SJ, Jones S, Moles D, Gill D. Psychosocial impact of hypodontia in children. *Am J Orthod Dentofac Orthop* 2010;137:35–41.
20. Cronbach I.J. Coefficient alpha and the internal structure of the tests. *Psychometrika* 1951;16:297–333.
21. Slade G.D., Spencer A.J. Development and evaluation of the oral health impact profile. *Community Dent Health* 1994;11:3–11.