Otitis Media and Caregiver Quality of Life: Psychometric Properties of the Modified Danish Version of the Caregiver Impact Questionnaire
Christian Hamilton Heidemann, Christian Godballe, Anette Drøhse Kjeldsen, Eva Charlotte Jung Johansen, Christian Emil Faber and Henrik Hein Lauridsen

Otolaryngology -- Head and Neck Surgery 2014 151: 142 originally published online 31 March 2014
DOI: 10.1177/0194599814528245

The online version of this article can be found at:
http://oto.sagepub.com/content/151/1/142

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
AMERICAN ACADEMY OF OTOLARYNGOLOGY--HEAD AND NECK SURGERY
F O U N D A T I O N
American Academy of Otolaryngology- Head and Neck Surgery

Additional services and information for Otolaryngology -- Head and Neck Surgery can be found at:

Email Alerts: http://oto.sagepub.com/cgi/alerts
Subscriptions: http://oto.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav

>> Version of Record - Jun 23, 2014
OnlineFirst Version of Record - Mar 31, 2014
What is This?
Otitis Media and Caregiver Quality of Life: Psychometric Properties of the Modified Danish Version of the Caregiver Impact Questionnaire

Christian Hamilton Heidemann, MD1,2, Christian Godballe, MD, PhD1,2, Anette Drøhse Kjeldsen, MD, PhD1,2, Eva Charlotte Jung Johansen, MD, PhD3, Christian Emil Faber, MD, PhD1,2, and Henrik Hein Lauridsen, MSc, PhD4

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

Abstract
Objective. Otitis media in children may have a considerable impact on caregiver quality of life. The disease-specific Caregiver Impact Questionnaire is designed to assess caregiver quality of life in relation to child otitis media. Assessment of the psychometric properties of this instrument is limited. This study assesses the psychometric properties of this instrument including validity, reproducibility, responsiveness, and interpretability.

Study Design. Longitudinal validation study.

Setting. Secondary care units.

Methods. Analyses were based on data from 435 families. Validity was assessed using confirmatory factor analysis, internal consistency, and hypothesis testing. Test–retest reliability and measures of smallest detectable change were investigated in the assessment of reproducibility. Responsiveness was investigated by means of hypothesis testing and receiver operating characteristic analysis. An anchor-based distribution method was applied for determining minimal important change as perceived by the respondent.

Results. Factor analysis confirmed the hypothesized 1-factor structure with an acceptable fit. Cronbach’s alpha was .90. In the analysis of construct validity, 88.9% of the hypothesized correlations were correctly predicted. Intraclass correlation coefficient was 0.87 and smallest detectable change corresponded to approximately one-fourth of the scale. Responsiveness was found to be good and a change score of 13.8 represented minimal important change.

Conclusion. The modified Danish version of the Caregiver Impact Questionnaire is a valid and reproducible measurement tool that is also sensitive to measuring change in the current setting. A change score representing minimal important change as perceived by the respondent is proposed. Results of this study support the use of this instrument.

Keywords
caregiver, CIQ, cross-cultural adaptation, factor analysis, minimal important change, otitis media, smallest detectable change, validation

Received December 19, 2013; revised February 10, 2014; accepted February 26, 2014.

Introduction
Otitis media (OM) is the leading cause of doctor consultations for preschool children,1 and studies have indicated that caregiver functioning and quality of life (QoL) may be negatively affected by OM.2–9 This is in concordance with studies on chronic childhood diseases such as asthma and diabetes.10,11 Insight into disease impact on caregivers, particularly in preschool children with chronic diseases or diseases like otitis media that may resemble a chronic disease for a period of time, is important for different reasons. Parent proxy reporting of the child health state is often

1Department of ENT Head & Neck Surgery, Odense University Hospital, Odense C, Denmark
2Institute of Clinical Research, Faculty of Health Sciences, University of Southern Denmark, Odense M, Denmark
3Ear Nose Throat Private Clinic, Odense C, Denmark
4Research Unit for Clinical Biomechanics, Institute of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense M, Denmark

Part of this article was presented at the 29th Politzer Society Meeting; November 15, 2013; Antalya, Turkey.

Corresponding Author:
Christian Hamilton Heidemann, MD, Department of ENT Head & Neck Surgery, Odense University Hospital, Sdr. Boulevard 29, 5000 Odense C, Denmark.
Email: chrheidemann@gmail.com
necesssary in children aged 4 years or younger due to limited cognitive and language abilities, and studies have indicated that caregivers’ own QoL may influence their proxy ratings of the child health state. Furthermore, parents play an important role in the cooperation between children and health care professionals. Quality of life of the parents may therefore influence the treatment strategy for the child. Therefore, reports on both child and caregiver QoL are important in assessing treatment outcomes and planning future treatment strategies.

The literature describing the impact of otitis media on caregiver QoL is limited. Most studies either included only a single item on questionnaires designed for assessment of child QoL or used a nonvalidated measure. Only 3 studies included validated disease-specific measures for assessing caregiver QoL as an individual construct. Two studies used the Family Functioning Questionnaire (FFQ) and a version of the FFQ designed for telephone surveys, respectively. The last study used the Caregiver Impact Questionnaire (CIQ), which is a further development of the FFQ.

Investigations on the psychometric properties of the CIQ are limited to 1 study assessing reliability and validity. The aim of this study is to translate the CIQ into Danish and assess the psychometric properties, with special attention given to the factor structure, sensitivity to measuring change, and the interpretability.

Patients and Methods

Patients and Design

Caregivers were enrolled from private Ear-Nose-Throat (ENT) clinics in southern Denmark as part of a study investigating the effects of ventilating tube treatment (VT) in children with recurrent acute otitis media (rAOM) and otitis media with effusion (OME). Inclusion criteria were (1) indication for VT established by an ENT specialist, (2) age 0 to 6 years, (3) no history of VT, and (4) the caregiver should be able to read, write, and understand Danish. Exclusion criteria were syndrome diseases, cleft lip and palate, or other concurrent illnesses with the potential to affect the QoL such as severe heart or lung disease. The study was approved by the Danish Data Protection Agency. According to the rules and regulations of the Danish Scientific Ethical Committee, approval is not required for studies that are purely observational.

Instruments and Procedure

The CIQ includes 6 items covering physical and emotional domains and an appertaining numerical rating scale (NRS-caregiver) for the assessment of global disease-specific QoL of the caregiver (see appendix, available at otojournal.org). The principles of forward and backward translations were applied in accordance with international guidelines. The few discrepancies between the original and back-translated versions were resolved at a consensus meeting (see appendix). In the original version, respondents were asked to recall symptom history pertaining to the previous 3 months. However, we chose a 4-week period to (1) streamline the recall period with the remaining questionnaires in the study and (2) reduce recall bias, which may be problematic in fluctuating diseases. Items were adjusted to scales of 0 to 100.

Outcome was measured at 3 time points. Caregivers completed questionnaires on the day the ENT specialist established indication for VT (pre-baseline), within 4 days prior to surgery (baseline) and at 1-month follow-up. The following questionnaires were included: pre-baseline (CIQ, NRS-caregiver, the Otitis Media-6 questionnaire [OM-6], and NRS-child), baseline, and follow-up (CIQ, NRS-caregiver, OM-6, NRS-child, relevant subscales of the 50-item Child Health Questionnaire [CHQ-PF50] and 36-item Short Form questionnaire [SF-36], and 3 questions regarding the number of interrupted nights, days absent from work or education, and number of times forced to cancel social activities because of OM in the child). Furthermore, a 7-point Global Perceived Effect (GPE) scale (Figure 1) matching the CIQ and NRS-caregiver was included at follow-up (a detailed description of all instruments is provided in the appendix). All pre-baseline questionnaires were handed out and completed on paper. For all subsequent questionnaires, caregivers were given the choice between paper-based questionnaires and electronic questionnaires. Eighty-two percent of caregivers completed subsequent questionnaires online. We regarded respondents who completed the baseline questionnaire more than 7 days after VT as not eligible for data analysis.

Statistical Analysis

Missing items were investigated at baseline and follow-up and managed in the following manner: if more than 50% of items were missing, the summary score was discarded.
Test–retest reliability was assessed by computing the intraclass correlation coefficient (ICC). An asymptotically distribution-free estimation method was applied as data were non-normally distributed. Model accuracy was based on the chi-square test and the following model fit indices: (1) comparative fit index (CFI), (2) root mean square error of approximation (RMSEA), and (3) standardized root mean square residual (SRMR). As the relatively large sample size has the potential to produce statistically significant chi-square values that are essentially unimportant, all significant chi-square values were interpreted in combination with the other fit indices. Model fit was interpreted as “acceptable” if CFI > 0.90, RMSEA < 0.08, and SRMR < 0.08. Model misSpecifications were investigated by calculating modification indices (MI). If the model required modifications, the MI along with content-related considerations were used. Last, internal consistency was assessed by calculating Cronbach’s alpha. Alpha should be between .70 and .95.

Structural validity. Confirmation of the hypothesized 1-factor structure is most adequately established with confirmatory factor analysis (CFA). An asymptotically distribution-free estimation method was applied as data were non-normally distributed. Model accuracy was based on the chi-square test and the following model fit indices: (1) comparative fit index (CFI), (2) root mean square error of approximation (RMSEA), and (3) standardized root mean square residual (SRMR). As the relatively large sample size has the potential to produce statistically significant chi-square values that are essentially unimportant, all significant chi-square values were interpreted in combination with the other fit indices. Model fit was interpreted as “acceptable” if CFI > 0.90, RMSEA < 0.08, and SRMR < 0.08. Model misSpecifications were investigated by calculating modification indices (MI). If the model required modifications, the MI along with content-related considerations were used. Last, internal consistency was assessed by calculating Cronbach’s alpha. Alpha should be between .70 and .95.

Construct validity. Hypotheses were constructed regarding correlations between items, between items and summary scores, and between summary scores of the different instruments. For example, scores on item 1 of the CIQ (lack of sleep) were expected to correlate positively and strongly (>0.5) with the number of interrupted nights because of OM, as an increase in the number of interrupted nights was expected to negatively influence the caregiver’s perception of lack of sleep. A higher percentage of correct predictions indicates stronger support for construct validity (see appendix).

Reproducibility. Assessment of reproducibility included analysis of test–retest reliability and smallest detectable change. Test–retest reliability was assessed by computing the intra-class correlation coefficient (ICC2.1.A). Criteria for inclusion in the test–retest analysis were (1) an interval of 2 to 14 days between pre-baseline and baseline measurements, (2) caregivers had to state that they perceived the state of OM in their children to be static between the repeated measurements, and (3) questionnaires at both measurements must be completed by the same caregiver. ICC of at least 0.70 is generally required as a minimum standard for test–retest reliability. The smallest detectable change (SDC) is based on the standard error of measurement (SEM), which is the variability in measurements (SD) of the same individual with a confidence of 95% and is expressed in the unit of the measurement. It was estimated by computing the square root of the within-subject variance of the respondents (SEM = \[\sqrt{\sigma_{\text{between}}^2 + \sigma_{\text{residual}}^2}\]). Variance components were obtained from a multilevel mixed effects model (restricted maximum likelihood estimates). Because SDC is the smallest amount of change in individuals that can be detected beyond measurement error with a confidence of 95%, it is calculated as SDC = 1.96*\[\sqrt{\text{SEM}}\].

Responsiveness. For assessment of criterion responsiveness, the GPE scale was included as an external anchor, and correlations between change scores and the GPE score were expected to be at least 0.5. Receiver operating characteristic (ROC) analyses were performed in which respondents were dichotomized in groups of “importantly improved” versus “stable” according to their responses on the GPE scale (Figure 1). Area under the curve (AUC) was interpreted as the probability of correctly identifying the “importantly improved” parents from “stable” parents. An AUC of 1.00 indicates perfect discrimination, whereas an AUC of 0.50 indicates discrimination no better than chance. Area under the curve should be at least 0.70. Construct responsiveness was assessed by hypothesizing that correlations between change scores of the instruments should be at least 0.5. Last, floor and ceiling effects were investigated at baseline as presence of these may hamper the possibility of detecting change. These were considered present if more than 15% achieved the highest or the lowest possible score.

Interpretability. Minimal important change as perceived by the respondent (MIC) was determined in the analysis of interpretability. An optimum cut-off point was retrieved from the ROC analysis and was used as a benchmark for the MIC score. By weighting sensitivity and specificity equally, the cut-off point was assumed to represent the lowest overall misclassification. The MIC was related to the SDC by computing the group size needed to achieve an SDC, that equals the MIC [N = (SDC/MIC)^2].

STATA® v. 12 IC (StataCorp) was used for all analyses.

Results

Four hundred ninety-one families were enrolled in the study. Fifty-six had to be excluded because of late baseline responses (>7 days postsurgery). Fifty-six percent of the children were male, median age was 1.46 years (interquartile range = 1.23), and 15% had rAOM, 47% had OME, and 38% had a mixed diagnosis (rAOM/OME). Response rates were 95% and 92% at baseline and follow-up, respectively. Missing items ranged from 0.0% to 1.1% for the 6 items. Scale and item descriptive data are presented in Table 1.

Validity. Confirmatory factor analysis was computed based on 397 valid responses and confirmed the 1-factor model. However, to obtain an acceptable fit, it was necessary to include correlations of 2 sets of error variances in the model. Modification indices revealed large covariances between items 5 and 6 and items 3 and 4. These covariances made conceptual sense, and correlations between error terms of the 2 item sets were included in the final model (Table 2). Standardized factor loadings ranged from 0.63 to 0.85 (Figure 2). The internal consistency equalled 0.90
Analysis of construct validity (N = 381) showed that 23 of the 27 (88.9%) hypothesized correlations were correctly predicted (see appendix).

Reproducibility. There was a mean of 6.7 days between test–retest measurements (N = 135). We found no significant differences between the test–retest subsample and the remaining study sample with regard to age, gender, diagnostic distribution, and baseline scores (see appendix). Test–retest reliability was acceptable (Table 3). Mean difference was close to zero, indicating no systematic difference between test–retest scores. Analysis of the SDC showed that only change scores exceeding 22.3 and 27.6 for the CIQ and NRS-caregiver, respectively, would be detectable beyond measurement error at the individual level.

Responsiveness. Change scores of the instruments correlated strongly with each other and the GPE score (N = 386) (Table 4). As anticipated, correlations between change scores of the disease-specific questionnaires (CIQ, OM-6, NRS-caregiver, and NRS-child) were higher than correlations with the generic questionnaires (CHQ-PF50 and SF-36; N = 389). Areas under the curve were 0.80 and 0.79 for the CIQ and NRS-caregiver, respectively, would be detectable beyond measurement error at the individual level.

Interpretability. The MIC was 13.8 points for the CIQ and 20.0 for the NRS-caregiver with a sensitivity and specificity of 0.7 or above. Smallest detectable change was larger than MIC for both instruments (Table 5), and the group size

Table 1. Item Scores and Summary Scores of the Caregiver Impact Questionnaire and Scores of the NRS-Caregiver at Baseline and Follow-Up.a

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>CIQ item 1</td>
<td>412</td>
<td>52.8</td>
</tr>
<tr>
<td>CIQ item 2</td>
<td>412</td>
<td>27.5</td>
</tr>
<tr>
<td>CIQ item 3</td>
<td>410</td>
<td>22.6</td>
</tr>
<tr>
<td>CIQ item 4</td>
<td>411</td>
<td>32.4</td>
</tr>
<tr>
<td>CIQ item 5</td>
<td>414</td>
<td>44.2</td>
</tr>
<tr>
<td>CIQ item 6</td>
<td>415</td>
<td>47.9</td>
</tr>
<tr>
<td>CIQ summary score</td>
<td>415</td>
<td>37.9</td>
</tr>
<tr>
<td>NRS-caregiver</td>
<td>415</td>
<td>53.9</td>
</tr>
</tbody>
</table>

Abbreviations: CIQ, Caregiver Impact Questionnaire; N, number of respondents; NRS, Numerical Rating Scale; SD, standard deviation.

aItem 1: lack of sleep; Item 2: absence from work or education; Item 3: canceling of family activities, such as trips, play dates, vacations; Item 4: changing daily activities, such as housework, shopping, or time with other siblings; Item 5: feeling nervous, agitated, or irritable; Item 6: feeling helpless or frustrated.

Table 2. Confirmatory Factor Analysis—Fit Statistics.

<table>
<thead>
<tr>
<th>Model (N = 397)</th>
<th>$\chi^2$</th>
<th>df</th>
<th>P Value</th>
<th>CFI</th>
<th>RMSEA</th>
<th>SRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original 1-factor structurea</td>
<td>88.257</td>
<td>9</td>
<td>&lt;.001</td>
<td>0.836</td>
<td>0.149</td>
<td>0.082</td>
</tr>
<tr>
<td>Item 5 with 6b</td>
<td>43.716</td>
<td>8</td>
<td>&lt;.001</td>
<td>0.926</td>
<td>0.107</td>
<td>0.049</td>
</tr>
<tr>
<td>Item 5 with 6 and 3 with 4c</td>
<td>23.599</td>
<td>7</td>
<td>.001</td>
<td>0.966</td>
<td>0.077</td>
<td>0.035</td>
</tr>
</tbody>
</table>

Abbreviations: CFI, comparative fit index; df, degrees of freedom; N, number of respondents; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual.

aOriginal structure without correlation of error terms.
bCorrelation of error terms of items 5 and 6 included in the model.
cCorrelation of error terms of items 5 and 6 and error terms of items 3 and 4 included in the model.

(N = 400). Analysis of construct validity (N = 381) showed that 23 of the 27 (88.9%) hypothesized correlations were correctly predicted (see appendix).

Figure 2. Path diagram displaying the final model with standardized factor loadings and error correlations.
required to detect a MIC beyond measurement error was 3 for the CIQ and 2 for the NRS-caregiver.

**Discussion**

The CIQ was found to be reliable and had good construct validity and responsiveness in the current setting. Analysis of the structural validity confirmed a 1-factor structure as expected from the literature. However, our results revealed item sets with a high proportion of shared variance, which may suggest overlap in item content. An acceptable fit was obtained in CFA when these covariances were allowed in the model. In addition, a change score of 13.8 on the CIQ and 20.0 on the NRS-caregiver represented an important change for the respondents but was within measurement error at the individual level. However, this change score will be beyond measurement error even in small study samples.

We used CFA to assess the structural validity of the CIQ. All 6 items loaded significantly and strongly on a single factor and an acceptable fit was obtained after allowing the error terms of 2 pairs of items to be correlated in the model. Large covariance was found especially between item

**Table 3. Reproducibility—Test–Retest Reliability and Smallest Detectable Change.**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean (SD) Difference</th>
<th>ICC (CI)</th>
<th>SEM</th>
<th>SDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIQ</td>
<td>135</td>
<td>–1.9 (11.3)</td>
<td>0.87 (0.82-0.91)</td>
<td>8.0</td>
<td>22.3</td>
</tr>
<tr>
<td>NRS-caregiver</td>
<td>135</td>
<td>–1.1 (14.1)</td>
<td>0.82 (0.76-0.87)</td>
<td>9.9</td>
<td>27.6</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; CIQ, Caregiver Impact Questionnaire; ICC, intraclass correlation coefficient; N, number of respondents; NRS, Numerical Rating Scale; SDC, smallest detectable change at the individual level; SEM, standard error of measurement.

*Difference between test–retest scores.

**Table 4. Responsiveness—Correlations between Change Scores of the Different Instruments and between Change Scores and the Score of the Global Perceived Effect Scale.*

<table>
<thead>
<tr>
<th></th>
<th>Mean (SD) Change</th>
<th>CHQ-PF50</th>
<th>SF-36</th>
<th>CIQ</th>
<th>NRS-Caregiver</th>
<th>OM-6</th>
<th>NRS-Child</th>
<th>GPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIQ</td>
<td>24.6 (22.3)</td>
<td>–0.69</td>
<td>–0.67</td>
<td>—</td>
<td>–0.78</td>
<td>0.72</td>
<td>–0.73</td>
<td>–0.55</td>
</tr>
<tr>
<td>NRS-caregiver</td>
<td>–27.9 (25.0)</td>
<td>0.59</td>
<td>0.60</td>
<td>–0.78</td>
<td>—</td>
<td>0.72</td>
<td>–0.73</td>
<td>0.81</td>
</tr>
</tbody>
</table>

Abbreviations: AUC, area under the curve; CHQ-PF50, 50-item Child Health Questionnaire; CIQ, Caregiver Impact Questionnaire; GPE, Global Perceived Effect scale; NRS, Numerical Rating Scale; OM-6, Otitis Media-6 questionnaire; SF-36, 36-item Short Form questionnaire.

*Spearman’s rho correlation coefficients.

*Change scores computed from summary scores of the included subscales.

![Diagram](CIQ.png)

**Figure 3.** Distribution of change scores after dichotomizing respondents in groups according to responses on the Global Perceived Effect scale. Optimum cut-off point retrieved from receiver operating characteristic analysis determined the minimal important change score (MIC).
5 (feeling nervous, agitated, or irritable) and item 6 (feeling helpless or frustrated). This makes conceptual sense as both items are aimed at assessing emotional domains and we believe the covariance can be explained by some degree of content overlap. Hence, it may be argued that 1 of these items should be omitted from the questionnaire. However, we recommend retaining both items in the questionnaire because omitting 1 of the 2 items covering emotional impact will leave this domain with very little impact on the summary score. However, acknowledging this issue as a possible weakness of the CIQ is important. Covariance was also found between item 3 (canceling of family activities, such as trips, play dates, vacations) and item 4 (changing daily activities, such as housework, shopping, or time with other siblings). Although there is an evident association between these items, this does not necessarily indicate content overlap. When OM in a child becomes a burden on family functioning, the family is likely to experience a negative impact on aspects covered by both items.

Our results support construct validity of the CIQ, and this is consistent with findings by the original developers.2 All correlations were 0.5 or above, with the highest correlations between summary scores of the disease-specific instruments (0.68-0.83).

Test–retest reliability was acceptable for both instruments. The original developers confirmed test–retest reliability by computing a Pearson’s correlation coefficient of 0.83. We used ICC because it takes systematic error into account.31 The measurement error as measured by the SDC was relatively large. Changes of less than 22% on the CIQ and 28% on the NRS-caregiver are not detectable beyond measurement error on an individual level. In comparison, Brouwer et al19 presented SEM values of 6.1 and 8.3 corresponding to SDC values of 17 and 23 for the FFQ and NRS-caregiver, respectively. The slightly poorer agreement found in our study may be explained by different factors. First, the inclusion of a more heterogeneous study population (children with both OME and rAOM) may have increased the error. Second, our administration method may have added a possible bias as the pre-baseline questionnaires were completed on paper in the ENT clinic and baseline questionnaires were completed electronically at home. Last, the differences between the test–retest scores may, in part, be explained by an intention to give socially desirable answers at the clinic or by being more distracted by external factors, for example, by a crying or impatient child.

Our findings showed that the CIQ is sensitive to measuring changes over time. Analysis of construct responsiveness revealed strong correlations between change scores of instruments measuring similar constructs. Furthermore, correlations between change scores of the CIQ and the GPE scores were what we expected. The AUC indicated that the CIQ was able to discriminate between “importantly improved” and “stable” caregivers. Last, no floor or ceiling effects were present, enabling bidirectional change scores.

When interpreting CIQ change scores, it is important to know whether results are statistically significant but also whether they are relevant for patients or clinicians. By including the GPE scale, we aimed at assessing the MIC as perceived by the respondent. However, to meaningfully interpret the importance of change scores, the MIC should be linked to measurement error (SDC). The ratio for the SDC with the MIC was smaller than 1, which implies that using the CIQ on an individual level is problematic. Respondents with change scores between 13.8 (MIC) and 22.3 (SDC) may have experienced important improvements but have a change score within measurement error. However, most clinical research is conducted to investigate changes at the group level and measurement error is decreased by \( \sqrt{N} \) in studies at the group level. In this situation, our results indicate that MIC will be beyond measurement error even in small studies. It should be noted that the group sizes represent minimal values and regular power calculations should be performed to assess actual sample sizes needed in studies.

### Methodological Considerations

The study population consisted of caregivers of children scheduled for VT, and disease severity may, to some extent, limit generalizability. Floor effect may be present in a population of less disease severity. Furthermore, the children in our study sample had a median age younger than 2 years, which is younger than in some studies.7,8,16,41,42 Language and cognitive development are limited at this age, which may negatively influence parent reporting, especially on the emotional domain. Second, a GPE scale was applied in the responsiveness analysis and for determining the MIC. Critical remarks have been made about such a transition measure with regard to its reliability and because it tends to depend more on the most recent measurement than on the first measurement (recall bias).43 Our analysis revealed the same issues (results not presented). However, we found the strongest correlations between change scores and the GPE score. Furthermore, we wanted to assess MIC as perceived by the respondent and not clinicians. To do so, it is necessary to include an anchor directly aimed at this issue, for example, a GPE scale.

### Table 5. The Minimal Important Change and Its Relation to the Smallest Detectable Change.

<table>
<thead>
<tr>
<th></th>
<th>Range</th>
<th>MIC</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Ratio MIC/SDC</th>
<th>Nstudy&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIQ</td>
<td>0-100</td>
<td>13.8</td>
<td>0.8</td>
<td>0.7</td>
<td>0.62</td>
<td>3</td>
</tr>
<tr>
<td>NRS-caregiver</td>
<td>0-100</td>
<td>20.0</td>
<td>0.8</td>
<td>0.7</td>
<td>0.73</td>
<td>2</td>
</tr>
</tbody>
</table>

Abbreviations: CIQ, Caregiver Impact Questionnaire; MIC, minimal important change; NRS, Numerical Rating Scale; SDC, smallest detectable change.

<sup>a</sup>Number of caregivers needed in a study for the scale to be able to detect MIC beyond measurement error.
Conclusion

The modified Danish version of the CIQ has proven to have good reproducibility, validity, and responsiveness. A change score of 13.8 represented respondent perceived minimal important change. This score is outside measurement error at the group level, implying that the CIQ is an accurate instrument to measure clinically important changes in clinical trials. Results of this study support the use of this instrument.

Acknowledgments

We want to thank the developers for permitting the use of the Caregiver Impact Questionnaire. We also want to thank the participating ENT specialists in private clinics and the participating families.

Author Contributions

Christian Hamilton Heidemann, wrote study protocol, questionnaire administration, data collection, data analysis, wrote manuscript, final approval; Christian Godballe, wrote study protocol, manuscript revision, final approval; Anette Drobhse Kjeldsen, contributed substantially to the conception and design of the study, manuscript revision, final approval; Eva Charlotte Jung Johansen, contributed substantially to the conception and design of the study, questionnaire administration and inclusion of patients, manuscript revision, final approval; Christian Emil Faber, contributed substantially to the conception and design of the study, manuscript revision, final approval; Henrik Hein Lauridsen, wrote study protocol, data analysis, manuscript revision, final approval.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: Research Foundation of the Region of Southern Denmark and the Foundation of Professional Development in Specialist Private Clinics. No role in study.

Supplemental Material

Additional supporting information may be found at www.otojournal.org/supplemental.

References

21. Nielsen S, Ruperto N, Herlin T, Pedersen FK. The Danish version of the Childhood Health Assessment Questionnaire...


