Development of otology specific outcome measure: Ear Outcome Survey-16 (EOS-16)

Juha T. Laakso [a, *], Juha Silvola [b], Timo Hirvonen [a], Samuli Suutarla [a, c], Ilkka Kivekäs [d], Riitta Saarinen [a], Lotta Haavisto [e], Jaakko Laitakari [f], Antti A. Aarnisalo [a], Aarno Dietz [g], Jussi Jero [e], Maija Hytönen [a], Saku T. Sinkkonen [a]

[a] Department of Otorhinolaryngology — Head and Neck Surgery, Head and Neck Center, Helsinki University Hospital and University of Helsinki, Helsinki, Finland
[b] Department of Ear, Nose, and Throat — Division of Surgery, Akershus University Hospital, Lørenskog, Norway
[c] Department of Otorhinolaryngology, Kymenlaakso Central Hospital, Kotka, Finland
[d] Department of Otorhinolaryngology — Head and Neck Surgery, Tampere University Hospital and Tampere University, Tampere, Finland
[e] Department of Otorhinolaryngology — Head and Neck Surgery, Turku University Hospital and University of Turku, Turku, Finland
[f] Department of Otorhinolaryngology — Head and Neck Surgery, Oulu University Hospital and PEDEGO Research Unit, University of Oulu and Medical Research Center Oulu, Finland
[g] Department of Otorhinolaryngology — Head and Neck Surgery, Kuopio University Hospital and University of Kuopio, Kuopio, Finland

ARTICLE INFO

ABSTRACT

Purpose: An important outcome measure of patient care is the impact on the patient's health-related quality of life (HRQoL). Current ear-specific HRQoL instruments are designed for one diagnosis and emphasize different subdivisions such as symptoms, hearing problems, psychosocial impact, and the need for care. The optimal length of the recall period has not been studied. For these reasons, a new survey is needed that would cover most chronic ear diseases.
Methods: A preliminary 24-item survey (EOS-24) was created. Untreated adult patients (included n = 186) with one of seven different chronic otologic conditions from all university hospitals in Finland were recruited to respond to EOS-24 and the 15D general HRQoL instrument. The recruiting otologists evaluated the severity of the disease and the disability caused by it. A control group was recruited. Based on the patients' responses in different diagnosis groups, the items were reduced according to pre-defined criteria. The resulting survey was validated using a thorough statistical analysis.
Results: The relevance and necessity of the original 24 items were thoroughly investigated, leading to the exclusion of 8 items and the modification of 1. The remaining 16 items were well-balanced between subdivisions and were useful in all seven diagnosis groups, thus constituting the final instrument, EOS-16. The most suitable recall period was three months.
Conclusions: EOS-16 has been created according to the HRQoL survey guidelines with a versatile nationwide patient population. The survey has been validated and can be used for a wide range of chronic ear diseases as a HRQoL instrument.

1. Introduction

Along with the patient's symptoms and findings, increasing attention has recently been paid to the subjective quality of life (QoL) (Litwin 2006; Lailach et al., 2018). For physicians, the interest is in restoring health-related QoL (HRQoL) that has been decreased by illness (D. L. Patrick and Bergner 1990; Cramer 2002). The World Health Organization defines HRQoL as a state of complete physical, mental, and social well-being and not merely the absence of disease (WHO 1948; Litwin 2006). In some fields of medicine, HRQoL instruments are used to monitor treatments and their effectiveness (Santana et al., 2015; Psotka et al., 2019; Sitlinger and Zafar 2018; Baumhauer 2017; Donovan et al., 2016). The HRQoL measure may

* Corresponding author. Department of Otorhinolaryngology — Head and Neck Surgery, Head and Neck Center, Helsinki University Hospital, P.O. Box 220, FI-00029, HUH, Helsinki, Finland.
E-mail address: juha.laakso@hus.fi (J.T. Laakso).

be generic or disease-specific. As opposed to the generic tool, the disease-specific HRQoL measure provides detailed information on the problems (Brittenden et al., 2019). The Sinonasal Outcome Test-22 (SNOT-22) (Hopkins et al., 2009) is a good example of a disease-specific HRQoL instrument used by clinicians worldwide in daily practice.

Previously published ear-specific HRQoL questionnaires are the Chronic Ear Survey (CES) (Nadol et al., 2000), Chronic Otitis Media Outcome Test (COMOT-15) (Baumann 2009), Chronic Otitis Media Questionnaire (COMQ-12) (Phillips et al., 2014), Zurich Chronic Middle Ear Inventory (ZCMEI-21) (Bachinger et al., 2016), Chronic Otitis Media Benefit Inventory (COMBI) (Phillips et al., 2017), and the Stapesplasty Outcome Test 25 (SPOT-25) (Lailach et al., 2017). The CES focuses on disease-specific symptoms and need for care (Nadol et al., 2000), whereas the other surveys also cover psychosocial aspects. The COMOT-15 strongly emphasizes hearing symptoms (7/15 items) (Baumann 2009). In the COMQ-12, there is only one item for psychosocial aspects, and the item reduction was not based on statistical analysis (Phillips et al., 2014). The ZCMEI-21 is a comprehensive ear-specific HRQoL survey (Bachinger et al., 2016) that emphasizes psychosocial aspects (8/21 items); most of its items have a recall period of two weeks, while the other inquiries emphasize symptoms over the last six months. Previous ear-specific surveys are limited to only chronic otitis media and one survey to otosclerosis only (SPOT-25). SPOT-25 has 25 items related to hearing and psychosocial aspects (Lailach et al., 2017). The COMBI serves only as a patient-reported outcome measure (PROM) because it is a better−worse type of inquiry (Phillips et al., 2017).

In our experience, the symptoms and challenges caused by different chronic ear diseases are rather similar but may differ in severity. Currently, there is no usable questionnaire in otology to cover a wide spectrum of chronic ear diseases. Thus, we decided to develop a novel ear survey with several specific qualities. We aimed to broaden the survey focus to include evaluation of most chronic ear diseases with a single instrument. The survey should be used to screen patients not only before treatment but also after treatment as a PROM. The survey should include fewer than twenty items to minimize the response burden and increase reliability (D. L. Patrick and Bergner 1990; Mullin et al., 2000); its text should be clear and unambiguous in order to apply to different dialects and socioeconomic backgrounds. The survey itself should be easy to navigate (Mullin et al., 2000). For high reproducibility and discrimination of answers, a combined visual and numerical scale, a modified Likert-type rating scale, is widely accepted (Aaronson 1989; Lund et al., 2005; Guyatt et al., 1993; McPhail and Haines 2010). The literature indicates that the recall period should be as short as possible but long enough to cover the natural course of the disease (Aaronson 1989; Topp et al., 2019; McPhail and Haines 2010). Keeping these specifications in mind, we established a core development team (CDT) to develop the questionnaire, and otologists participating from different university hospitals enabled nationwide patient recruitment.

2. Materials and methods

This multicenter study was designed to develop a HRQoL instrument for chronic ear diseases. The study design is provided in Fig. 1.

2.1. Ethics

The study protocol was approved by the Helsinki University Hospital Ethics Committee (§49/17), and a research permit was obtained for each research center. Informed consent was obtained from all participants.

2.2. Patient demographics and diagnostic groups

Altogether, 263 patients were recruited by experienced specialists from ENT outpatient clinics in all five university hospitals in Finland. The patients were recruited between June 2017 and December 2018, and they had one of the following chronic ear diseases: secretory otitis media (SOM), tympanic membrane perforation with dry middle ear, chronic otitis media without cholesteatoma (COM), cholesteatoma, otosclerosis, exostosis, or chronic otitis externa (Fig. 1). The distribution of the patients into different diagnostic groups represents their prevalence in the recruiting outpatient clinics (Table 1). The inclusion criteria were no prior ear surgery (unless an earlier tympanostomy did not resolve the problem), sufficient Finnish language skills (native or comparable), and at least 18 years of age. All patients fulfilling the inclusion criteria were recruited to the study regardless of their disease severity in order to obtain a representative study sample. The study population represents an unselected group to be examined and treated at outpatient clinics in Finland including those with diverse educational backgrounds, differences in job description, and wealth. From the 263 participants, 77 were excluded for missing answers (n = 68), a missing specialist evaluation (n = 8), or withdrawal (n = 1) (Fig. 1). In total, 186 participants were included. The otosclerosis group was roughly twice as big as the other diagnostic groups (Table 1; n = 65), and thus, it was overrepresented. Therefore, it was split in half by recruitment order for the upcoming analysis. The first half (n = 32) was used to select the recall period and item reduction, and the second (n = 33) was used in the validation. The item reduction was thus conducted on responses from 153 patients. A control group comprised of 23 Helsinki University Hospital staff members who, according to their own reports, had no ear problems. Patient characteristics and the distribution of diagnoses are shown in Table 1. In total, 174 of 186 participants responded to the 15D.

2.3. EOS-24 creation

The core development team (CDT) examined the CES, COMQ-12, COMOT-15, and ZCMEI-21 questionnaires for potential items (questions or claims). During examination, redundant items were removed, items were edited if needed, and new items were created. In the end, 24 potential items covering four subdivisions (ear symptoms, hearing impairment, psychosocial impact, and need for care) were selected and translated into Finnish. The items were formulated into a questionnaire inspired by the SNOT-22 (Hopkins et al., 2009). This novel survey was named the Ear Outcome Survey-24 (EOS-24, Table 2). For the sake of clarity and unambiguity, physicians, patients, and hospital staff evaluated the text and form on several occasions, which led to changes. There were eventually twelve development versions of the EOS-24, the last of which was used to collect data.

For every 24 items, the patients simultaneously evaluated their symptoms and the effect of the disease during the previous two weeks, three months, and six months. Recall periods of two weeks and six months were chosen according to previous surveys, and the CDT added an additional recall period of three months. Answers were presented using a modified 6-pointLikert-type scale where "DCM" corresponded to "Doesn't concern me"; 0 to "I have this problem but no physical or emotional impact"; 1 to "mild problem"; 2 to "moderate problem"; 3 to "severe problem"; and 4 to "very severe problem."

The patients also responded to a generic HRQoL survey, the 15D (Sintonen and Pekurinen 1993), in which the symptoms are scored with a logarithmic formula from very severe (0) to no impact (1). The recruiting ENT specialist made a simultaneous assessment of
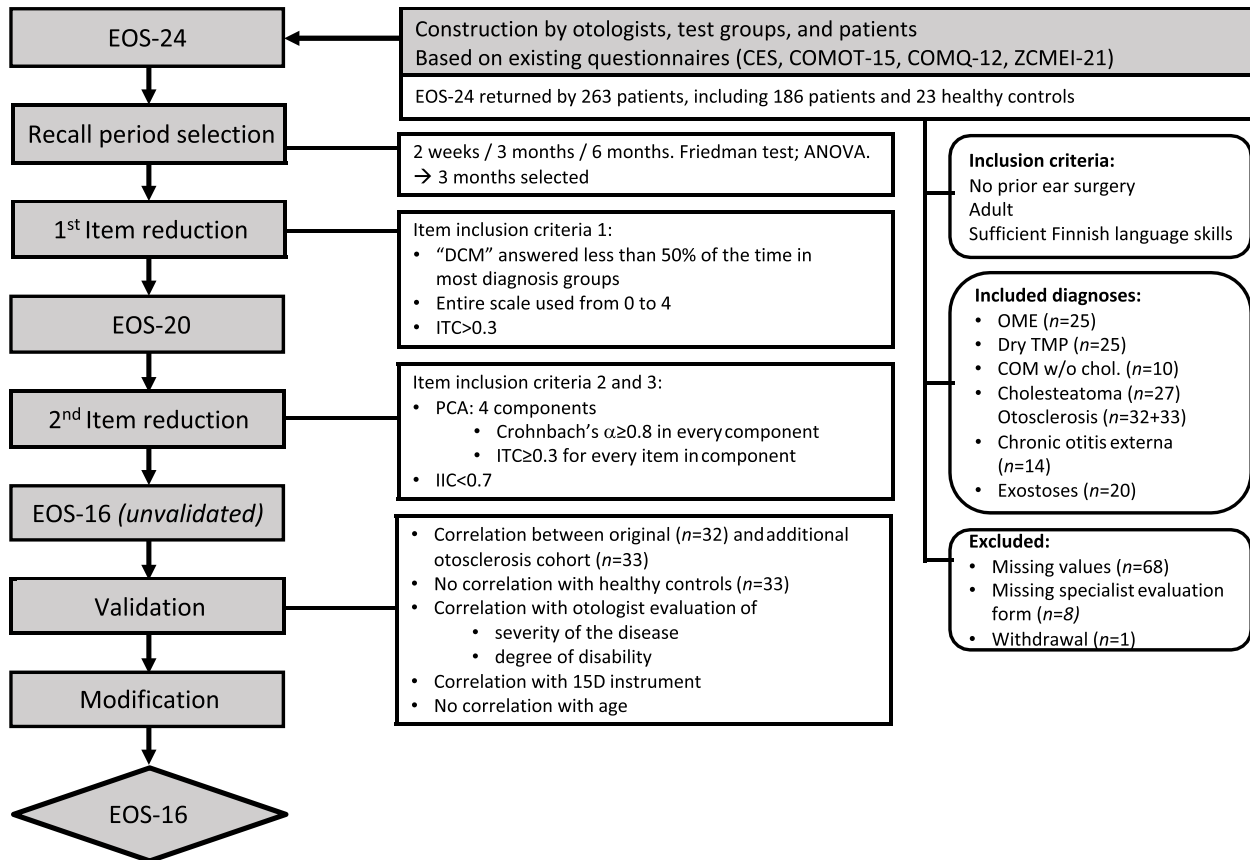
Fig. 1. Study design. OME, Middle ear with effusion; Dry TM perforation, dry tympanic membrane perforation; COM w/o chol., chronic otitis media without cholesteatoma.

Table 1
Patient characteristics and diagnosis. OME, Middle ear with effusion; Dry TM perforation, dry tympanic membrane perforation; COM w/o chol., chronic otitis media without cholesteatoma.

	Item reduction group ($n = 153$)	Otosclerosis group for item reduction ($n = 32$)	Otosclerosis group for validation ($n = 33$)	Control group of healthy individuals ($n = 23$)
Age (y), mean ± SEM (range)	52 ± 17 (18–88)	47 ± 14 (18–67)	50 ± 10 (28–71)	39 ± 8 (27–55)
Sex (male), n (%)	87 (57)	17 (53)	14 (42)	10 (42)
Diagnosis				
OME, n (%)	25 (16)	0	0	0
Dry TM perforation, n (%)	25 (16)	0	0	0
COM w/o chol., n (%)	10 (7)	0	0	0
Cholesteatoma, n (%)	27 (18)	0	0	0
Otosclerosis, n (%)	32 (21)	32 (100)	33 (100)	0
Chronic otitis externa, n (%)	14 (9)	0	0	0
Exostoses, n (%)	20 (13)	0	0	0

the severity of the disease and the degree of disability caused by it using a Likert-type scale from 0 to 4 (0 no impact; 1 very mild impact; 2 mild impact; 3 moderate impact; 4 severe impact).

2.4. Item reduction

Item reduction was based on statistical analysis. To select the most suitable recall period, we first compared the mean values between recall periods (Supplemental Table, Appendix 1). Our study group was slightly right-skewed, and transformation with square root was used to achieve a normal distribution before further analysis between recall periods. The following analyses

were performed with a repeated measures ANOVA to detect the statistical difference between the recall periods.

All 24 items were then evaluated using the following inclusion criteria: (1) the entire scale was used from "Doesn't concern me" (DCM) or 0 to 4; (2) the percentage of "DCM" was calculated for all diagnostic groups and did not exceed 50% in most diagnostic groups. The items were divided into groups with principal component analysis. Internal consistency was evaluated, and the inclusion criteria were (3) item-total correlation (ITC)>0.3, meaning at least a moderate correlation; and (4) Cronbach's $\alpha > 0.8$, meaning a high level of consistency within an item group. To evaluate which items correlate too well and are considered

Table 2
EOS-24 presented with data for recall period selection and item reduction. Table contains all diagnoses (n = 153).

EOS-24 Original items (24 pc.)	Item group	Comparison of Recall periods			Recall period of 3 months			
		2 weeks mean (SD)	3 months mean (SD)	6 months mean (SD)	Min	Max	% N (Q1–Q21) or 0 (Q22–24)	ITC
1. I have had pain in my ear.	Ear symptoms	0.65 (0.98)	0.95 (1.12)	1.12 (1.24)	0	4	33.7	0.527
2. I have had an itch in my ear.	Ear symptoms	0.99 (1.20)	1.02 (1.21)	1.09 (1.26)	0	4	27.0	0.541
3. I have felt pressure in my ear.	Ear symptoms	0.87 (1.16)	1.00 (1.12)	1.03 (1.14)	0	4	31.3	0.632
4. I have felt moisture in my ear.	Ear symptoms	0.86 (1.22)	0.97 (1.19)	1.03 (1.29)	0	4	31.9	0.543
5. I have had discharge in my ear.	Ear symptoms	0.63 (1.15)	0.75 (1.14)	0.90 (1.30)	0	4	42.3	0.453
6. I have heard a buzzing or ringing sound in my ear (tinnitus).	Hearing	1.20 (1.16)	1.31 (1.14)	1.38 (1.15)	0	4	17.8	0.431
7. I have experienced dizziness or disequilibrium.	Balance[a]	0.53 (0.88)	0.66 (0.95)	0.69 (1.02)	0	4	38.0	0.343
8. I have experienced vertigo.	Balance[a]	0.25 (0.67)	0.24 (0.60)	0.30 (0.74)	0	3[b]	63.2[c]	0.160[d]
9. My hearing has worsened.	Hearing	1.50 (1.31)	1.64 (1.32)	1.62 (1.29)	0	4	17.8	0.639
10. I have had problems hearing because of background noise.	Hearing	1.67 (1.39)	1.75 (1.37)	1.71 (1.34)	0	4	16.6	0.624
11. I have had difficulty locating the direction of a sound.	Hearing	1.31 (1.39)	1.40 (1.38)	1.32 (1.33)	0	4	26.4	0.651
12. I have been using a hearing aid due to hearing loss. I find this challenging.	Hearing	0.14 (0.61)	0.13 (0.59)	0.12 (0.53)	0	4	78.5[c]	0.233[d]
13. I have had difficulty hearing a phone or an alarm clock.	Hearing	0.70 (1.10)	0.73 (1.11)	0.71 (1.09)	0	4	47.9[c]	0.555
14. I have had trouble in group discussions because of my hearing.	Hearing	1.18 (1.29)	1.25 (1.28)	1.16 (1.24)	0	4	25.2	0.654
15. Protecting my ear from water has restricted my life. [If you do not protect your ear from water, select "DCM."]	Need for care[a]	0.63 (1.14)	0.64 (1.14)	0.65 (1.13)	0	4	49.1[c]	0.179[d]
16. Because of my ear problems, I do not socially interact as much with other people.	Quality of life	0.80 (1.12)	0.86 (1.15)	0.78 (1.07)	0	4	37.4	0.654
17. Because of my ear problems, I face challenges in my daily activities/at school/at work.	Quality of life	0.92 (1.19)	1.04 (1.20)	1.03 (1.16)	0	4	31.9	0.667
18. Symptoms related to my ears limit my hobbies.	Quality of life	0.59 (1.04)	0.69 (1.09)	0.68 (1.10)	0	4	40.5	0.502
19. I fear that the symptoms related to my ears will worsen in the future.	Quality of life	1.44 (1.37)	1.53 (1.34)	1.48 (1.35)	0	4	17.2	0.581
20. My ear problems make me sad.	Quality of life	0.92 (1.16)	1.02 (1.22)	1.03 (1.23)	0	4	31.3	0.607
21. I feel like my ear problems have negatively affected my quality of life.	Quality of life	1.10 (1.15)	1.24 (1.15)	1.19 (1.16)	0	4	23.3	0.709
22. I have consulted a doctor because of my ear problems.	Need for care[a]	0.76 (0.89)	1.40 (1.22)	1.86 (1.34)	0	4	26.4	0.304
23. I have used oral antibiotics because of my ear problems.	Need for care[a]	0.06 (0.33)	0.25 (0.70)	0.45 (0.93)	0	4	76.1[c]	0.324
24. I have used antibiotic ear drops (as prescribed).	Need for care[a]	0.35 (0.68)	0.69 (1.10)	1.08 (1.36)	0	4	58.3[c]	0.366

[a] Group includes Balance and Need for care
[b] whole scale was not used from 0 to 4 (Q8)
[c] % "DCM" ("Doesn't concern me") and 0 exceeds 50 in most diagnostic groups (Q8, Q12, Q13, Q23)
[d] ITC is below 0.3 in most diagnosis groups (Q8).

redundant, Spearman's rank correlation was used. (5) Correlation coefficients $r_s > 0.7$ were considered too high. A high correlation coefficient between items meant that one should replace all of these items. (6) We aimed for an even distribution of ear symptoms, hearing problems, need for treatment, and psychosocial effects across subgroups.

2.5. Validation

The Pearson correlation coefficient (PCC) was used to evaluate the final EOS instrument, the EOS-16, with the 15D and age. Spearman's rank correlation was used to evaluate dependencies between EOS-16 scores and physicians' evaluations on disease severity and disability caused by the disease. Analyses of EOS-16 scores between the otosclerosis groups of item reduction and the additional validation group were based on the Mann–Whitney U Test, as were the EOS-16 score comparison between the reduction group and the control group.

2.6. Statistics

Statistical analyses were performed using IBM SPSS Statistics for Macintosh version 24 (Armonk, NY). Graphs were made using Prism 8 (GraphPad, San Diego, CA). Forms missing more than 3 answers out of 72 were excluded. Occasional missing answers were replaced with medians of that particular item from the same responder. The significance level was set at $p \leq 0.05$. A professional statistician confirmed the validity of the methods used.

3. Results

3.1. Recall period selection

Based on the EOS-24 responses, the different recall periods were compared. They demonstrated a cumulative increase in the score means from 20.0 ± 1.2 (mean $\pm$ SEM) for 2 weeks, to 23.2 ± 1.2 for three months, and 24.4 ± 1.3 for six months ($p < 0.001$, Supplemental Table, Appendix 1). The two-week mean was statistically lower than the three-month ($p < 0.001$; repeated measures ANOVA) or six-month mean ($p < 0.001$), suggesting that patients did not have as many symptoms during the previous 2 weeks as during the longer periods. Between three months and six months, there was no statistically significant difference ($p = 0.169$). The same analyses were also performed within individual diagnostic

groups (Supplemental Tables, Appendices 1–7). The SOM and otosclerosis groups had no statistical difference between two weeks, three months, and six months. All the other diagnostic groups demonstrated fewer symptoms in the prior two weeks than in the longer periods. In general, a shorter recall period increases the accuracy of the survey (Aaronson 1989; McPhail and Haines 2010; Topp et al., 2019), and there was no difference between three and six months; therefore, the CDT selected 3 months as the recall period for the survey (Appendix 9).

3.2. First item reduction

The reduction was initiated by examining the first three criteria (Table 2). The whole scale (DCM to 4) was not used in item number 8 (Q8), which led to its exclusion. For most diagnoses, the percentage of "DCM" was over 50% in Q12, Q13, and Q23, leading to their exclusion. Q24 had 58% "DCM" responses overall, but it was a good item in the COM group (DCM 30%, Supplemental Table, Appendix 3) and an excellent item in the chronic external otitis group (DCM 0%, Supplemental Table, Appendix 6). The CDT's assessment was that this item was essential for these two diagnostic groups and should not be removed. As expected, Q15 had a low item-total correlation in the otosclerosis group, but it was below expectations as well in the chronic external otitis, tympanic membrane perforation, and COM groups. The CDT's assessment was that the item was elongated and unclear, which placed its effectiveness and item-total correlation significantly below expectations. However, the item was not removed since its content was considered important, and thus, the item was modified (Fig. 2). For this modification, we held two cognitive debriefings (n = 10 responders) to make this item unambiguous and easy to read. These analyses were also performed within individual diagnostic groups with virtually similar results (Supplemental Tables, Appendices 1–7).

3.3. Second item reduction

The principal component analysis divided the items into four components, where component 1 encompassed ear symptoms and most of the need for care items. Component 2 tracked hearing items and balance, component 3 encompassed all psychosocial items, and component 4 only contained one need for care item, Q15. Component 1 (Q1, Q2, Q3, Q4, Q5, Q22, and Q24) had ITC>0.3 for every item and Cronbach's α = 0.841. Between Q4 and Q5, the correlation coefficient (r_s = 0.75, $p < 0.001$) was too high, which led to the exclusion of Q5 due to a smaller mean value (0.97 versus 0.75). Component 2 (Q6, Q7, Q9, Q10, Q11, and Q14) had ITC>0.3 for every item and Cronbach's α = 0.863. Between Q10 and Q14, the correlation was too high (r_s = 0.78, $p < 0.001$), which led to the exclusion of Q14 due to a smaller mean value (1.75 versus 1.25). Q9 and Q10 also shared a strong correlation (r_s = 0.78, $p < 0.001$), but neither was removed since Q9 was considered a filter item for later use in an electronic form. Component 3 (Q16, Q17, Q18, Q19, Q20, and Q21) had ITC>0.3 for every item and Cronbach's α = 0.877. There were similar types of items, but their internal correlations were not as strong as suspected. Based on our sixth criteria, two items had to be excluded. The two highest correlation coefficients were between the pairs Q16, Q17 (r_s = 0.64, $p < 0.001$) and Q20, Q21 (r_s = 0.63, $p < 0.001$); from these pairs, the lowest means were used as a reason to exclude Q16 (0.86 versus 1.04) and Q20 (1.02 versus 1.24). The item reduction led to the creation of the EOS-16 (Cronbach's alfa = 0.87, Fig. 2 and Supplemental Fig., Appendix 8). When analyzing the different diagnostic groups separately, the results were similar, suggesting that we had selected the most appropriate items (Supplemental Tables, Appendices 1–7).

3.4. Validation

To test whether item reduction led to a change in the nature of the survey, we compared EOS-16 scores between the item reduction otosclerosis group (n = 32) and the control otosclerosis group (n = 33) and found no difference (Fig. 3A, p = 0.207, Mann–Whitney U Test). The EOS-16 clearly differentiated patients from controls (Fig. 3B, $p < 0.001$, Mann–Whitney U Test). There was a correlation between the EOS-16 and 15D scores (Fig. 3C, r = −0.36, $p < 0.001$). Also, the EOS-16 score correlated with the severity of the disease (Fig. 3D, r_s = 0.36, $p < 0.001$) and the degree of disability caused by it (Fig. 3E, r_s = 0.40, $p < 0.001$), as assessed by otologists. The EOS-16 scores had no correlation with the patients' age (Fig. 3F, r = 0.053, p = 0.512), suggesting that the survey does not measure the patients' age and fragility but rather the harm caused by the disease.

3.5. Modifications after validation

The response scale used in the EOS-24 was a 6-pointLikert-type scale. In 77% of responses without a perceived adverse ear problem, "DCM" was chosen as the response rather than option 0. We combined these two groups in the EOS-16 because they seemed to measure the same construct (Fig. 2).

4. Discussion

In modern healthcare, more emphasis is given to treatments that are proven effective. Together with objective measures, an important outcome is the subjective HRQoL. To justify different treatments and compare their efficacy, we need valid HRQoL measures. Although previously published ear-specific HRQoL inquiries are available, none of them covers a wide variety of chronic otologic conditions in a single questionnaire (Nadol et al., 2000; Phillips et al. 2014, 2017; Baumann 2009; Bachinger et al., 2016; Lailach et al., 2017; Vlastos et al., 2009). Therefore, there was a need for a broad-spectrum otology-specific HRQoL instrument developed in a rigorous statistical process (Fig. 1) and validated according to COSMIN guidelines (Mokkink et al., 2010; Terwee et al., 2018). In addition, several otology-specific factors, as well as different diagnoses together and separately, were considered to make the questionnaire valuable in clinical practice.

We believe the recall period is an important specification when measuring HRQoL, but it has not been considered in great detail in most of the previous otologic outcome measures (Nadol et al., 2000; Phillips et al. 2014, 2017; Baumann 2009; Bachinger et al., 2016; Lailach et al., 2017; Vlastos et al., 2009). The recall period should be the shortest possible, accounting for the natural course of the disease (Aaronson 1989; Topp et al., 2019; McPhail and Haines 2010). Therefore, we chose to study non-treated patients.

Adequate coverage of the items is limited by the respondent's motivation (D. L. Patrick and Bergner 1990; Mullin et al., 2000). We started with 24 items but were able to reduce the number to 16 after thorough statistical analyses. We had patients participating nationwide from each university hospital district, representing different dialects and cultural backgrounds to ensure that the scores reflect the entire Finnish population.

Another important aspect is a balance between different otologic HRQoL subdivisions: ear symptoms, hearing impairment, psychosocial impact, and need for care. In previous inquiries, the emphasis of the different subdivisions varies (Nadol et al., 2000; Phillips et al. 2014, 2017; Baumann 2009; Bachinger et al., 2016; Lailach et al., 2017; Vlastos et al., 2009). In the EOS-16, the items are evenly distributed into four subdivisions, which makes it versatile and sensitive to several dimensions of various otologic conditions.

Ear Outcome Survey-16 (EOS-16©)					v1.1

Name:			Date:		

Personal ID:					

Below is a list of the symptoms and problems associated with ear disorders. With the help of the survey, we will find their occurrence and impact on your Quality of Life.

Select the level of difficulty that best describes your problems over <u>THE LAST THREE MONTHS</u>. If you have symptoms in both ears, answer the questions about the ear with more severe symptoms. If you answer the survey for a post-treatment control visit, please answer for the treated ear.	No problem	Mild problem	Moderate problem	Severe problem	Very severe problem
1. I have had pain in my ear.	0	1	2	3	4
2. I have had an itch in my ear.	0	1	2	3	4
3. I have felt pressure in my ear.	0	1	2	3	4
4. I have felt moisture in my ear.	0	1	2	3	4
5. I have heard a buzzing or ringing sound in my ear (tinnitus).	0	1	2	3	4
6. I have experienced dizziness or disequilibrium.	0	1	2	3	4
7. My hearing has worsened.	0	1	2	3	4
8. I have had problems hearing because of background noise.	0	1	2	3	4
9. I have had difficulty locating the direction of a sound.	0	1	2	3	4
10. Protecting my ear from water has restricted my life.	0	1	2	3	4
11. Because of ear problems, I face challenges in my daily activities / at school / at work.	0	1	2	3	4
12. Symptoms related to my ears limit my hobbies.	0	1	2	3	4
13. I fear that the symptoms related to my ears will worsen in the future.	0	1	2	3	4
14. I feel like my ear problems have negatively affected my quality of life.	0	1	2	3	4

Select the option that best describes your situation. Answer the questions for the <u>PREVIOUS THREE MONTHS</u>.	I haven't	Once	Twice	3-4 times	Over 4 times
15. I have consulted a doctor because of my ear problems.	0	1	2	3	4
16. I have used antibiotic ear drops (as prescribed).	0	1	2	3	4

Overall score: _____

Fig. 2. The non-validated English version of the Ear Outcome Survey-16 (EOS-16©).

For validation, we compared EOS-16 scores between the otosclerosis patients of the item reduction group and the validation group and found no difference (Fig. 3A), confirming that no item with a significant effect on the score was removed during the item reduction. There was a clear-cut difference between the healthy controls and the patients, as expected (Fig. 3B). A positive correlation was found between the EOS-16 and 15D scores, reflecting the EOS-16's ability to measure the general HRQoL to some extent (Fig. 3C). Importantly, a correlation was seen between the EOS-16 score and the physician's evaluation of the severity of the disease (Fig. 3D) and the degree of disability caused by it (Fig. 3E),

demonstrating the EOS-16's applicability to clinical practice. It is important to note that at this stage, the EOS-16 has only been validated as a HRQoL measure whose goal is to develop the EOS-16 into a PROM in the future with pre- and post-intervention measures for different diagnoses.

The study has some limitations. The suitability of the EOS-16 for adolescents warrants future studies. At the moment, the EOS-16 has been validated in Finnish (Supplemental Fig., Appendix 8), but validations have been planned for other languages. When comparing the number of patients in our study with other ear survey reports (Nadol et al., 2000; Baumann 2009; Vlastos et al.,
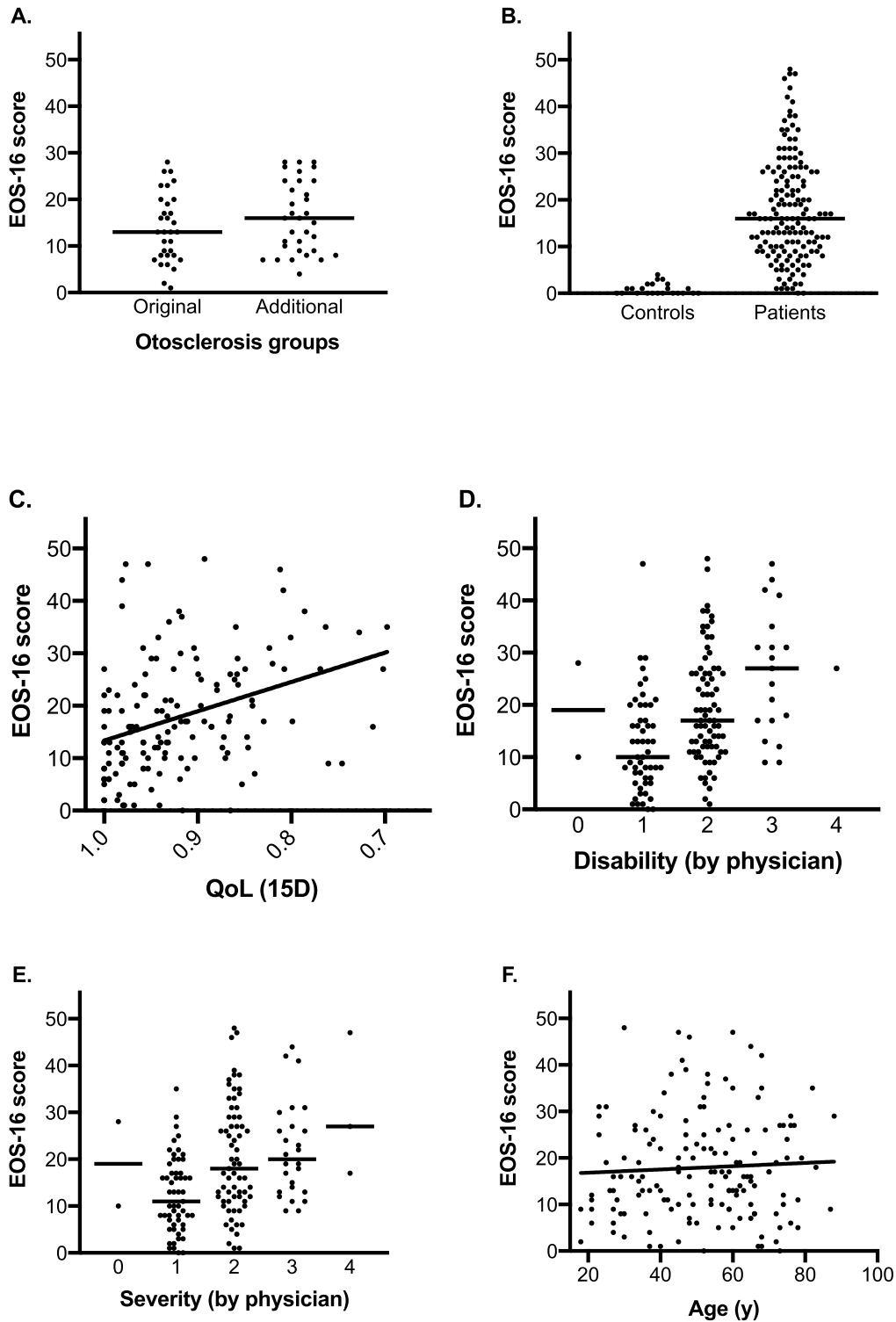
Fig. 3. Validation process. A Comparison of EOS-16 scores between otosclerosis patients in the item reduction group and additional group ($p = 0.207$). **B** EOS-16 score comparison between subjects and healthy controls ($p<.001$). **C** Association between the 15D and EOS-16 ($r = -0.36$, $p<.001$) demonstrates (positive) correlation. **D** EOS-16 scores and degree of disability demonstrates correlation ($r_s = 0.40$, $p<.001$). **E** Correlation between the EOS-16 and the severity of disease ($r_s = 0.36$, $p<.001$). **F** Correlation between EOS-16 scores and patients' age was insignificant ($r = 0.053$, $p = 0.512$). Horizontal lines represent averages in the graphs.

2009; Phillips et al. 2014, 2017; Bachinger et al., 2016; Lailach et al., 2017), we have significantly more patients, although the number of patients in each diagnostic group could still be larger. Due to the number of patients, we were unable to use the Rasch analysis (Boone 2016), but this is taken into account in further development.

HRQoL tool development is an ongoing process, and the EOS-16 is currently used as a validated HRQoL instrument to gather information on otologic patients' symptoms and to help in everyday clinical work-up. In the near future, it will be used as a PROM to study the effectiveness and cost-effectiveness of treatments in and between different chronic ear diseases (Caulley et al., 2019; Kruk et al., 2018; Neumann and Sanders 2017).

5. Conclusion

The EOS-16 was created according to the HRQoL survey guidelines with a versatile nationwide patient population. The survey is validated and can be used for a wide range of chronic ear diseases as a HRQoL instrument in everyday otologic practice.

Declaration of competing interest

None.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.joto.2021.01.003.

References

Aaronson, Neil K., 1989. Quality of life assessment in clinical trials: methodologic issues. Contr. Clin. Trials 10, 195–208.

Bachinger, D., Roosli, C., Ditzen, B., et al., 2016. 'Development and validation of the Zurich chronic middle ear inventory (ZCMEI-21): an electronic questionnaire for assessing quality of life in patients with chronic otitis media. Eur. Arch. Oto-Rhino-Laryngol. 273, 3073–3081.

Baumann, I., Kurpiers, B., Plinkert, P., et al., 2009. Entwicklung und validierung des chronic otitis Media Outcome Test 15 (COMOT-15). HNO Springer-Verlag.

Baumhauer, Judith F., 2017. 'Patient-reported outcomes — are they living up to their potential? N. Engl. J. Med. 377, 6–9.

Boone, William J., 2016. 'Rasch analysis for instrument development: why, when, and how? CBE-Life Sci. Educ. 15, rm4.

Brittenden, J., Cooper, D., Dimitrova, M., et al., 2019. 'Five-year outcomes of a randomized trial of treatments for varicose veins. N. Engl. J. Med. 381, 912–922.

Caulley, Lisa, Hunink, Myriam G., Kilty, Shaun, et al., 2019. 'Evidence-based medicine in otolaryngology part 9: valuing health outcomes. Otolaryngology-Head Neck Surg. (Tokyo) 160, 11–21.

Cramer, J.A., 2002. Principles of health-related quality of life: assessment in clinical trials. Epilepsia 43, 1084–1095.

Donovan, J.L., Hamdy, F.C., Lane, J.A., et al., 2016. 'Patient-reported outcomes after monitoring, surgery, or radiotherapy for prostate cancer. N. Engl. J. Med. 375, 1425–1437.

Guyatt, Gordon H., Feeny, David H., Patrick, Donald L., 1993. 'Measuring health-related quality of life. Ann. Intern. Med. 118, 622–629.

Hopkins, C., Gillett, S., Slack, R., et al., 2009. 'Psychometric validity of the 22-item sinonasal outcome test. Clin. Otolaryngol. 34, 447–454.

Kruk, M.E., Gage, A.D., Arsenault, C., et al., 2018. 'High-quality health systems in the sustainable development goals era: time for a revolution. Lancet Global Health 6, e1196–e1252.

Lailach, S., Schenke, T., Baumann, I., et al., 2017. [Development and validation of the stapesplasty outcome test 25 (SPOT-25)]. HNO 65, 973–980.

Lailach, S., Baumann, I., Zahnert, T., et al., 2018. '[State of the art of quality-of-life measurement in patients with chronic otitis media and conductive hearing loss]. HNO 66, 578–589.

Litwin, M.S., 2006. Health-related quality of life. In: Penson, D.F., Wei, J.T. (Eds.), Clinical Research Methods for Surgeons.

Lund, I., Lundeberg, T., Sandberg, L., et al., 2005. 'Lack of interchangeability between visual analogue and verbal rating pain scales: a cross sectional description of pain etiology groups. BMC Med. Res. Methodol. 5, 31.

McPhail, S., Haines, T., 2010. 'Response shift, recall bias and their effect on measuring change in health-related quality of life amongst older hospital patients. Health Qual. Life Outcome 8, 65.

Mokkink, L.B., Terwee, C.B., Patrick, D.L., et al., 2010. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual. Life Res.: Int. J. Qual. Life.Aspect.Treat.Care.Rehabil. 19, 539–549.

Mullin, P.A., Lohr, K.N., Bresnahan, B.W., et al., 2000. 'Applying cognitive design principles to formatting HRQOL instruments. Qual. Life Res.: Int. J. Qual. Life.-Aspect.Treat.Care.Rehabil. 9, 13–27.

Nadol Jr., J.B., Staecker, H., Gliklich, R.E., 2000. 'Outcomes assessment for chronic otitis media: the Chronic Ear Survey. Laryngoscope 110, 32–35.

Neumann, Peter J., Sanders, Gillian D., 2017. 'Cost-Effectiveness analysis 2.0. N. Engl. J. Med. 376, 203–205.

Patrick, D.L., Bergner, M., 1990. 'Measurement of health status in the 1990s. Annu. Rev. Publ. Health 11, 165–183.

Phillips, J.S., Haggard, M., Yung, M., 2014. 'A new health-related quality of life measure for active chronic otitis media (COMQ-12): development and initial validation. Otol. Neurotol. 35, 454–458.

Phillips, J.S., Haggard, M., Spencer, H., et al., 2017. The Chronic Otitis Media Benefit Inventory (COMBI): development and validation of a dynamic quality of life questionnaire for chronic ear disease. Otol. Neurotol. 38, 701–707.

Psotka, M.A., Fiuzat, M., Carson, P.E., et al., 2019. 'Design of a "lean" case report form for heart failure therapeutic development. JACC (J. Am. Coll. Cardiol.): Heart Fail. 7, 913–921.

Santana, Maria J., Haverman, Lotte, Kate Absolom, et al., 2015. 'Training clinicians in how to use patient-reported outcome measures in routine clinical practice. Qual. Life Res. 24, 1707–1718.

Sintonen, Harri, Pekurinen, Markku, 1993. 'A fifteen-dimensional measure of health-related quality of life (15D) and its applications. In: Quality of Life Assessment: Key Issues in the 1990s. Springer.

Sitlinger, A., Zafar, S.Y., 2018. 'Health-related quality of life: the impact on morbidity and mortality. Surg. Oncol. Clin. 27, 675–684.

Terwee, C.B., Prinsen, C.A.C., Chiarotto, A., et al., 2018. 'COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual. Life Res.: Int. J. Qual. Life.Aspect.Treat.Care.Rehabil. 27, 1159–1170.

Topp, J., Andrees, V., Heesen, C., et al., 2019. 'Recall of health-related quality of life: how does memory affect the SF-6D in patients with psoriasis or multiple sclerosis? A prospective observational study in Germany. BMJ Open 9, e032859.

Vlastos, I.M., Kandiloros, D., Manolopoulos, L., et al., 2009. 'Quality of life in children with chronic suppurative otitis media with or without cholesteatoma. Int. J. Pediatr. Otorhinolaryngol. 73, 363–369.

Who, 1948. Constitution of the World Health Organization, Basic Documents. WHO, Geneva.