International Journal of Audiology

ISSN: (Print) (Online) Journal homepage: informahealthcare.com/journals/iija20

Psychometric properties of the Swedish version of the international outcome inventory – alternative interventions (IOI-AI) – ear surgery (IOI-AI^{op})

Ylva Dahlin Redfors, Andreas Björsne & Caterina Finizia

To cite this article: Ylva Dahlin Redfors, Andreas Björsne & Caterina Finizia (08 Apr 2024): Psychometric properties of the Swedish version of the international outcome inventory – alternative interventions (IOI-AI) – ear surgery (IOI-AI^{op}), International Journal of Audiology, DOI: 10.1080/14992027.2024.2332774

To link to this article: https://doi.org/10.1080/14992027.2024.2332774

INTERNATIONAL JOURNAL OF AUDIOLOGY
https://doi.org/10.1080/14992027.2024.2332774

ORIGINAL ARTICLE

OPEN ACCESS Check for updates

Psychometric properties of the Swedish version of the international outcome inventory – alternative interventions (IOI-AI) – ear surgery (IOI-AI^{op})

Ylva Dahlin Redfors[a,b], Andreas Björsne[b,c,d] and Caterina Finizia[a,b]

[a]Department of Otorhinolaryngology, Head and Neck Surgery, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; [b]Department of Otorhinolaryngology, Head and Neck Surgery, Region Västra Götaland, Sahlgrenska University Hospital, Gothenburg, Sweden; [c]Unit of Audiology, Department of Health and Rehabilitation, Institute of Neuroscience and Physiology, University of Gothenburg, Gothenburg, Sweden; [d]Hearing Organization, Habilitation & Health, Region Västra Götaland, Gothenburg, Sweden

ABSTRACT

Objective: The aims of this study were to adapt the Swedish version of the International Outcome Inventory for Hearing Aids (IOI-HA) to the International Outcome Inventory for Alternative Interventions (IOI-AI) in the context of ear surgery (IOI-AI^{op}) and to test the psychometric properties.

Design: The validated Swedish questionnaire IOI-HA was adapted to the IOI-AI^{op} by omitting the question about hearing aid use and changing the term "hearing aid" to "surgery" in the remaining items. The validity, component structure and reliability of the IOI-AI^{op} were assessed.

Study sample: Subjects diagnosed with otosclerosis and undergoing stapedotomy were included in the study (n = 162).

Results: High mean scores were noted for all items. Ceiling effects were noted, most pronounced for the satisfaction item. Principal component analysis (PCA) yielded a two-component structure explaining 77.5% of the variance. The test-retest reliability measured by intra class correlation coefficient was >0.9, and the internal consistency coefficient measured by Cronbach's alfa was >0.8.

Conclusion: The IOI-AI^{op} showed good psychometric properties. However, ceiling effects were observed. The two-component solution was in line with previous factor analyses of the IOI-HA and the IOI-AI. The comprehensive IOI-AI^{op} is recommended as a useful tool to evaluate patient perspectives after ear surgery.

ARTICLE HISTORY
Received 19 April 2023
Revised 4 March 2024
Accepted 11 March 2024

KEYWORDS
IOI-HA; outcome measures; self-report; hearing impairment; psychometric properties; reliability; otosclerosis; stapedotomy

Introduction

Hearing is one of the most important senses for communicating and interacting with other people. Another important feature of hearing is that it gives us information about our surroundings. Losing the ability to hear or experiencing uncorrected hearing loss can have a substantial impact on everyday life, leading to difficulties in conversations, social isolation, lower quality of life, depression and an increased risk of developing dementia (Arlinger 2003).

To assess hearing and aspects related to hearing disability, audiometry must be accompanied by patient-related outcome measures (PROMs). PROMs can assess the subjective impact of hearing loss as well as the outcome of hearing rehabilitation, regardless of whether the intervention is hearing aid acquisition or ear surgery. A vast majority of existing questionnaires and studies are related to pure sensorineural hearing loss. In mixed or conductive hearing loss, there are sometimes other treatment options, such as middle ear surgery, middle ear implants and bone-anchored hearing aids, where fewer PROMs are available.

In systematic reviews analysing outcomes after treatment for mixed or conductive hearing loss the authors concluded that there is an underuse of PROMs. Only 22% of the assessed articles contained a PROM and only 11% in middle ear surgery (Hill-Feltham et al. 2021; Ostevik et al. 2021). In mixed and conductive hearing loss the most frequently used questionnaires were the Abbreviated Profile of Hearing Aid Benefit (APHAB) and Glasgow Benefit Inventory (GBI). In ear surgery, the most frequently used questionnaires were the APHAB and Speech, Spatial and Quality of Sounds (SSQ) (Cox and Alexander 1995; Gatehouse and Noble 2004; Robinson, Gatehouse, and Browning 1996). It is important that an assessment tool covers key areas of importance for hearing health (Granberg et al. 2014). The APHAB and SSQ lack items related to psychological issues, while the GBI is an intervention-specific questionnaire in otorhinolaryngology and lacks items related to hearing. To date, there is no golden standard. A questionnaire that could be used between different populations and interventions to facilitate comparison, especially in areas of conductive or mixed hearing loss with treatment modality options is needed. The International Outcome Inventory for Hearing Aids (IOI-HA), is a comprehensive 7-item questionnaire initially developed to assess hearing aid outcomes in a research context (Cox and Alexander 2002; Cox,

CONTACT Ylva Dahlin Redfors ylva.dahlin-redfors@gu.se Department of Otorhinolaryngology, Sahlgrenska University Hospital, Gröna Stråket 9, Gothenburg, SE-413 45, Sweden

Supplemental data for this article can be accessed online at https://doi.org/10.1080/14992027.2024.2332774.

Alexander, and Beyer 2003). The IOI-HA has been validated and has been shown to have good psychometric properties. Furthermore, the IOI-HA has been translated into many languages, including Swedish (Brännström and Wennerström 2010; Öberg, Lunner, and Andersson 2007). The questionnaire has been recommended as an assessment tool to compare intervention outcomes. It is short, with only a few items covering key areas. Furthermore, it is not dependent on cultural activities or context (Granberg et al. 2014). IOI-HA has been used both in the context of hearing aid acquisition and in the area of implantable hearing aids such as bone-anchored hearing aids, middle ear implants and cochlear implants (Arlinger, Nordqvist, and Öberg 2017; Heggdal et al. 2021; Zahnert et al. 2016).

A version for non-hearing aid-based interventions has also been developed, the International Outcome Inventory – Alternative Interventions (IOI-AI). Three possible areas of use were proposed (1) hearing assistive technology, (2) training and/or counselling, and (3) aural surgery (Noble 2002). The IOI-AI has above all been used to evaluate communication programs and has shown good psychometric properties (Hickson, Worrall, and Scarinci 2006; Kramer et al. 2005). For the IOI-AI aural surgery, a six-item version was suggested, with the question about hearing aid use omitted (Noble 2002). To our knowledge, the IOI-AI in the context of ear surgery (IOI-AI^{op}) has not been commonly used. The comprehensive IOI-AI^{op} could be a useful tool to assess patients' perspectives in the context of The Swedish National Quality Register for Otosclerosis Surgery since a validated patient reported outcome measures is missing.

The aim of this study was to analyse the psychometric properties of the IOI-AI^{op} by testing its validity and reliability in a group of otosclerosis subjects who underwent stapedotomy.

Methods

Adaptation of the questionnaire

The IOI-AI^{op} was developed by adapting the Swedish version of the IOI-HA to the aural surgery context. Question 1 (hearing aid use) was omitted, and in the following six questions, the word "hearing aid" was changed to "surgery" (Noble 2002).

Validation of the Swedish version of the IOI-AI^{op}

Study population

The study population consisted of two cohorts of subjects with hearing loss due to otosclerosis.

The first cohort was part of a larger prospective study. These subjects were included prospectively at two university clinics during 2017–2021 prior to surgical intervention (stapedotomy). The subjects were invited to participate in the study when visiting the clinic and were included after given informed consent. The inclusion criteria were as follows: (1) ages 20–65 years, (2) healthy without severe chronic health issues, and (3) air conduction (AC) PTA_4 (mean of frequencies 0.5, 1, 2 and 4 kHz) $\geq$ 30 dB HL and bone conduction (BC) $PTA_4 \leq$ 40 dB HL. Air bone gap (ABG) at 0.5 and 1 kHz $>$ 20 dB. 4) Good knowledge of the Swedish language.

The subjects were divided into two groups based on their prior experience with hearing aid use.

The second cohort was recruited from the Swedish quality register for otosclerosis surgery 1–2 years after stapedotomy. The inclusion criteria were as follows: (1) registered in the quality register for otosclerosis surgery, (2) stapedotomy performed in 2017–2019 and (3) had one-year follow-up. The subjects were invited to participate in the study with oral and written information. After giving informed consent, they were included in the study.

The included groups were as follows: Group STp) "primary stapedotomy" – stapedotomy without prior hearing aid use or prior ear surgery, included from the prospective group; Group STs) "secondary stapedotomy" – stapedotomy after prior hearing aid use, included from the prospective group; Group STq) stapedotomy included from the Swedish quality register for otosclerosis. Sixty seven percent of the subjects in the STq group had hearing aid experience.

The subjects completed the IOI-AI^{op}, Glasgow Benefit Inventory (GBI), Glasgow Hearing Aid Benefit Profile (GHABP), Short Form Health Survey 36 (SF-36) and Hospital Anxiety and Depression Scale (HADS) at 6 and 12 months after the intervention. The cohort from the quality register completed the questionnaires 12–24 months after the intervention. The questionnaires were sent via e-mail, and the responses were collected in an analysis platform (esMaker®, Entergate). Subjects who did not answer the questionnaire in 2–3 weeks were given one reminder. For test-retest reliability measures, 21 subjects answered the questionnaire a second time within 3 weeks of completing the first questionnaire. The subjects who answered the test-retest questionnaires were chosen to represent the whole cohort in age (mean age 47.5 years), sex (67% women) and inclusion group (43% from the quality register).

Questionnaires

IOI-AI^{op}

The IOI-AI^{op} was developed from the seven-item questionnaire IOI-HA by omitting the first item about hearing aid use and changing the word "hearing aid" to "surgery" in all remaining items. Hence, the IOI-AI^{op} comprises six items (1) benefits, (2) residual activity limitations (RAL), (3) satisfaction, (4) residual participation restriction (RPR) (5) impact on others and (6) quality of life (Noble 2002). Each item is scored on a Likert scale ranging from 1 to 5, where 5 represents the most favourable outcome. Factor analyses of the IOI-HA and IOI-AI have revealed a two-factor structure: Factor 1 (questions 1, 2, 4 and 7) represents hearing aid satisfaction, and Factor 2 (questions 3, 5 and 6) represents participation restrictions, meaning that the responses can be categorised and described by two factors (Cox and Alexander 2002; Kramer et al. 2002). Factors are calculated by adding the scores from the included items. Since item 1(use) was omitted in the IOI-AI^{op}, only items 2, 4 and 7 (benefit, satisfaction and QoL) were included in what is referred to as Factor 1^{op} in the analysis. Thus, the possible range of scores for both Factor 1^{op} and Factor 2^{op} were 3–15 (Cox and Alexander 2002).

GBI. The GBI is a generic questionnaire that is administered after interventions in the field of otorhinolaryngology. The questionnaire was developed by Robinson, Gatehouse, and Browning (1996). The questionnaire includes 18 items that are divided into three subscales: general health (12 items), social support (3 items) and physical health (3 items). The items are answered using a Likert scale ranging from 1 to 5. The subscales are then transformed to scores ranging from −100 to +100. Zero represents no change, −100 the worst scenario and +100 the best. It was recently translated and validated in Swedish (Redfors et al. 2019).

GHABP. The GHABP was developed by Gatehouse in 1999 to be used in hearing aid rehabilitation (Gatehouse 1999). The

questionnaire measures initial and residual disability, initial handicap and hearing aid benefit, satisfaction and use in predefined listening situations. The predefined listening situations are as follows: (1) listening to the television when the volume is adjusted for others, (2) having a conversation with one person in quiet, (3) having a conversation on a busy street or in a shop, and (4) having a conversation with several people in a group. The items are answered using a five-graded Likert scale (1 to 5). The subscales (initial and residual disability, initial handicap, hearing aid benefit, satisfaction and use) are developed by calculating the mean value of the different listening situations. A higher score for the disability and handicap subscales indicate a higher degree of disability and handicap. While, a higher score for the use, satisfaction and benefit subscales indicate a more favourable outcome. The questionnaire has recently been translated to Swedish and has been shown to have good psychometric properties (Dahlin Redfors, Jönsson, and Finizia 2022).

SF-36. The SF-36 is a generic questionnaire assessing health-related quality of life (Gandek et al. 2004; Stewart 1992). The questionnaire encompasses both physical and mental health aspects across 36 items and 8 subscales, including physical functioning (FP), role limitations due to physical problems (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role limitations due to emotional problems (RE), and mental health (MH). A score is calculated for each subscale. The score range between 0 (worse) and 100 (best). Two summary scores are then calculated; the physical component summary (PCS) and the mental component summary (MCS) (Stewart 1992). The questionnaire has been validated and translated into Swedish and has been shown to have good psychometric properties comparable with the original studies (Sullivan and Karlsson 1998; Sullivan, Karlsson, and Ware 1995; Taft, Karlsson, and Sullivan 2004).

HADS. The HADS is a frequently used questionnaire both in the research context and in the clinical setting. The HADS assesses signs of depression and anxiety across 14 items (Zigmond and Snaith 1983). The items are answered using a four-graded Likert scale (0 to 3). The scores for each item are added to a maximal score of 21 for each of the two subscales anxiety and depression. A score greater than 10 indicates an emotional problem (HADS-A >10 probable anxiety HADS-D >10 probable depression) (Zigmond and Snaith 1983). Different cut-off limits have been proposed depending on clinical setting, studied population and the desired level of sensitivity and specificity (Wu et al., 2021). The questionnaire has been validated and translated into Swedish and has shown good psychometric properties (Lisspers, Nygren, and Söderman 1997).

Audiometry

Pure tone audiometry for air and bone conduction (AC and BC) was performed in line with the ISO-standard 8253-1:2010 according to nationwide clinical routine, i.e. hearing thresholds were obtained using the ascending method, and contralateral masking was applied when indicated by the result. The frequencies of 0.5, 1, 2, 4, 6, and 8 kHz were measured for AC, and frequencies of 0.5, 1, 2, and 4 kHz were measured for BC. Pure tone averages for frequencies of 0.5, 1, 2, and 4 kHz (PTA^4) were calculated for AC and BC as well as for air bone-gap (ABG). The current study included measurements performed prior to and at 1 year after surgery (range 0.5–2 years).

Statistical analyses

Mean values, ranges and standard deviations were calculated for descriptive statistics.

Kruskal-Wallis and the chi-square tests were used to test differences between STp, STs and STq overall, regarding age and sex. Post-hoc testing between groups regarding age were performed with Fisher's permutation test and no corrections for multiple tests were made.

Statistical significance was set at $p < 0.05$, and all tests were two-sided. IBM SPSS Statistics for Windows, Version 27.0. (Armonk, NY:IBM Corp.) was used.

Floor and ceiling effects were examined and were considered present if >15% of the subjects achieved the highest or lowest scores (Terwee et al. 2007).

Principal component analysis (PCA) with Varimax rotation and Kaiser normalisation was used to explore underlying component structure of the IOI-AI^{op} items (*components* are the PCA counterpart to the *factors* extracted from a factor analysis). Components with an eigenvalue >1 were extracted.

Convergent and discriminant validity was assessed by performing Spearman's correlation analysis to obtain the correlations between IOI-AI^{op} subscales and the subscales of the other questionnaires included in the study. Hypotheses were made prior to the study regarding whether the subscales of the included questionnaires had a high correlation, predicting convergent validity or lower correlations supporting discriminant validity. The hypotheses are presented in Supplement 1 accessible at http://tandfonline.com/doi/suppl. According to Terwee et al., positive ratings are present if >75% of the predefined hypotheses are in accordance with the performed correlation analysis (Terwee et al. 2007).

Criterion validity was assessed by examining the correlations of factors with pure tone audiometry prior to and 1 year after intervention using Spearman's correlation analysis. Correlation coefficients ≤0.39 were regarded as weak, coefficients 0.4–0.59 as moderate and coefficients ≥0.6 as strong (Cohen 1988).

Test-retest reliability was assessed by calculating the intraclass correlation coefficient (ICC). Reliability and internal consistency were evaluated by calculating Cronbach's alpha. To indicate the strength of agreement, ≤ 0.20 was regarded as poor, 0.21–0.40 as slight, 0.41–0.60 as moderate, 0.61–0.80 as good, and 0.81–1.00 as very high (Fayers 2007).

Results

One hundred sixty-two subjects were included in the study. Sixty-four percent of the study population was women, and the mean age at intervention was 49.7 (±13.4) years. Statistically significant difference was observed between all groups regarding age (STp versus STs $p = 0.050$, STp versus STq $p < 0.001$, STs versus STq $p = 0.026$). Subjects in the primary stapedotomy group had the lowest mean age (43.7 ± 11.0 years) while subjects in the quality register group had the highest mean age (54.2 ± 14.1 years). Included in the validity calculations were subjects ($n = 148$) who had answered the 12–24 months questionnaires.

The hearing loss in the intervention ears was purely conductive or mixed and of moderate severity. Subjects in group STp had predominantly unilateral hearing loss, while groups STs and STq had predominantly bilateral hearing loss. Demographic data and hearing levels are presented in Tables 1 and 2.

The questionnaires were completed 6–24 months after the intervention.

Table 1. Demographic data.

	All n = 162	STp n = 47	STs n = 29	STq n = 86	p value (chi2, degrees of freedom)	p value (Kruskal-Wallis H, degrees of freedom)
Sex (% female, (n female))	64% (103)	64% (30)	76% (22)	59% (51)	p = 0.28 (2.57, 2)	
Age at intervention Years (±SD)	49.7 (13.4)	43.7 (11.0)	47.7 (8.9)	54.2 (14.1)		p < 0.001 (22.21, 2)

Groups: (STp) Stapedotomy, without prior hearing aid use or prior ear surgery – primary intervention; (STs) Stapedotomy, with prior hearing aid use – secondary intervention; (STq) Stapedotomy, included from the Swedish quality register for otosclerosis surgery-quality register.
*For comparisons between groups, the nonparametric Kruskal-Wallis and the chi-square tests were used. Fisher's permutation test was used for univariate analyses. Statistically significant differences were observed between all groups regarding age.
(STp versus STs $p = 0.050$, STp versus STq $p < 0.001$, STs versus STq $p = 0.026$).

Table 2. Pure tone audiometry.

Groups	All n = 148		STp n = 38		STs n = 24		STq n = 85	
	ST	CL	ST	CL	ST	CL	ST	CL
Pre intervention								
PTA_4 AC, dB HL (±SD)	54 (15.1)	26 (18.3)	47 (9.2)	15 (14.5)	54 (10.1)	28 (17.7)	56 (17.2)	30 (18.2)
10^{th} percentile	38	5	35	1	38	5	41	11
50^{th} percentile	51	21	48	9	52	22	54	26
90^{th} percentile	71	51	58	35	69	55	84	56
PTA_4 BC, dB HL (±SD)	24 (12.1)	19 (12.8)	19 (6.7)	11 (7.1)	20 (8.7)	17 (10.2)	28 (13.4)	22 (14.0)
10^{th} percentile	11	5	10	4	8	5	12	8
50^{th} percentile	21	16	18	10	20	15	28	19
90^{th} percentile	41	39	29	21	28	28	46	41
ABG, dB (±SD)	29 (9.5)	10 (11.2)	27 (7.3)	6 (10.8)	33 (9.0)	11 (12.7)	29 (10.3)	11 (10.7)
10^{th} percentile	17	−1	16	−2	17	−1	17	1
50^{th} percentile	28	6	27	2	32	9	27	8
90^{th} percentile	41	24	36	15	43	26	46	23
Post intervention								
PTA_4 AC, dB HL (±SD)	31 (14.1)	26 (18.1)	25 (7.7)	14 (12.8)	27 (10,1)	28 (18.7)	35 (15.7)	30 (17.8)
10^{th} percentile	19	5	16	2	19	8	19	9
50^{th} percentile	29	24	25	10	25	25	31	30
90^{th} percentile	49	50	35	34	42	53	59	54
PTA_4 BC, dB HL (±SD)	21 (12.8)	19 (13.3)	15 (7.6)	11 (5.9)	18 (9.4)	16 (10.6)	25 (14.0)	24 (14.8)
10^{th} percentile	9	5	5	2	6	4	10	8
50^{th} percentile	18	15	15	10	16	14	22	20
90^{th} percentile	38	36	26	18	26	30	45	48
ABG, dB (±SD)	10 (6.8)	9(10.9)	10 (7.8)	4 (10.6)	9 (4.3)	12 (14.4)	10 (7.0)	10 (8.8)
10^{th} percentile	3	−1	3	−4	4	−1	3	1
50^{th} percentile	9	6	9	1	8	6	8	9
90^{th} percentile	17	24	17	15	15	33	17	23

Groups: (STp) Stapedotomy, without prior hearing aid use or prior ear surgery – primary intervention; (STs) Stapedotomy, with prior hearing aid use – secondary intervention; (STq) Stapedotomy, included from the Swedish quality register for otosclerosis surgery – quality register.
ST = stapedotomy, intervention ear; CL = contralateral ear; PTA_4 mean of frequencies 0.5, 1, 2, and 4 kHz; dB = decibel; HL = hearing level; AC = air conduction; BC = bone conduction; ABG = air-bone gap.
Mean values, standard deviations (SD) and percentiles (10^{th}, 50^{th} and 90^{th}) are presented.

Table 3. Item level descriptive data of the IOI-AI^{op}.

Items	All n = 162	STp n = 47	STs n = 29	STq n = 86
Benefit	3.94 ± 1.35(1–5)	3.79 ± 1.30(1–5)	4.38 ± 1.18(1–5)	3.88 ± 1.4(1–5)
RAL	3.68 ± 1.18(1–5)	3.68 ± 1.02(1–5)	3.66 ± 1.14(1–5)	3.68 ± 1.29(1–5)
Satisfaction	4.4 ± 1.16(1–5)	4.26 ± 1.17(1–5)	4.62 ± 0.90(2–5)	4.40 ± 1.23(1–5)
RPR	4.43 ± 0.89(1–5)	4.38 ± 0.71(2–5)	4.34 ± 0.90(2–5)	4.49 ± 0.97(1–5)
Impact on others	4.33 ± 0.94(1–5)	4.40 ± 0.9(1–5)	4.41 ± 0.82(3–5)	4.26 ± 1.0(1–5)
Quality of life	3.64 ± 1.17(1–5)	3.30 ± 1.05(2–5)	4.10 ± 1.14(2–5)	3.66 ± 1.19(1–5)
Factor 1^{op}	12.01 ± 3.16(3–15)	11.43 ± 3.05(4–15)	13.10 ± 2.53(6–15)	11.95 ± 3.34(3–15)
Factor 2^{op}	12.43 ± 2.60(3–15)	12.47 ± 1.97(6–15)	12.41 ± 2.53(7–15)	12.40 ± 2.94(3–25)

Groups: (STp) Stapedotomy, without prior hearing aid use or prior ear surgery – primary intervention; (STs) Stapedotomy, with prior hearing aid use – secondary intervention; (STq) Stapedotomy, included from the Swedish quality register for otosclerosis surgery – quality register.
RAL = residual activity limitation; RPR = residual participation restriction.
Factor 1^{op}: benefit, satisfaction, QoL; Factor 2^{op}: RAL, RPR, impact on others.
Mean values are presented with standard deviation (SD) and range.
Each item is scored on a Likert scale ranging from 1 to 5, where 5 represents the most favourable outcome.

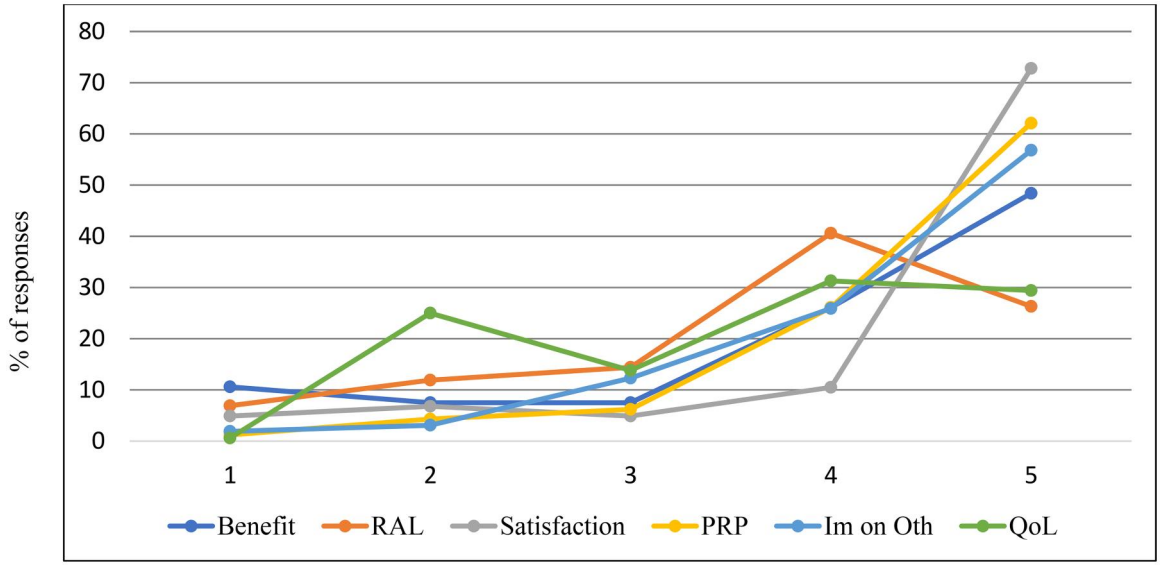

Figure 1. Distribution of the IOI-AI^{op} responses.
RAL = residual activity limitation; RPR = residual participation restriction; Im on Oth = impact on others; QoL = quality of life.
Each item is scored on a Likert scale ranging from 1 to 5, where 5 represents the most favourable outcome.

Table 4. Principal component analysis of the IOI-AI^{op}.

	Factor 1^{op}	Factor 2^{op}
Benefit	**0.88**	0.31
RAL	**0.62**	**0.61**
Satisfaction	**0.82**	0.18
RPR	0.11	**0.91**
Impact on Others	0.26	**0.88**
QoL	**0.81**	0.09

Principal component analysis was performed with Varimax rotation and Kaiser normalisation.
Components with an eigenvalue > 1 was extracted.
RAL = residual activity limitation; RPR = participation restriction; QoL = quality of life.

Validity

Item- and subscale level statistics are presented in Table 3. The distribution of responses is presented in Figure 1. Floor effects were not detected; however, ceiling effects were.

Residual activity limitation and participation restriction did not differ between the groups despite differences in uni- and bilateral hearing loss. Group STs, which included subjects with bilateral hearing loss and previous hearing aid experience, reported higher levels of satisfaction, benefit and quality of life than the other groups.

Construct validity

The Kaiser-Meyer-Olkin measures of sampling adequacy (KMO) test was 0.790, and the Bartlett's test of sphericity was significant with a p value <0.001, indicating that the data were appropriate for PCA. Principal component analysis yielded two components with eigenvalue > 1, explaining 77.5% of the variance. Items with component loadings > 0.80 corresponding to Factor 1^{op} included items assessing benefit, satisfaction and quality of life. Items with components loadings >0.80 corresponding to Factor 2^{op} were items assessing residual participation restriction and impact on others. The component loadings for residual activity limitation were similar to those of Factor 1^{op} (0.62) and 2^{op} (0.61) (Table 4).

Convergent and discriminant validity were calculated by examining the correlations between IOI-AI^{op} subscales and the subscales of the GBI, GHABP, SF36 and HADS. Hypotheses were developed prior to the analysis and were based on the concept being measured, see Supplement 1. Eighty-five percent of the hypotheses were supported by the correlation analysis, thus fulfilling the criteria of >75% agreement (Terwee et al. 2007). The strongest correlations were observed between the hearing-specific questionnaire GHABP and items regarding benefit and satisfaction, as could be expected (rho= +0.749, $p < 0.001$ and rho= +0.702, $p < 0.001$ respectively). A strong correlation was also detected between GHABP, item residual disability and Factor 2^{op} (rho= +0.780, $p < 0.001$) representing participation restrictions. Weak correlations were observed between the IOI-AI^{op} factors and the SF-36 as well as the social functioning and physical functioning subscales of the GBI (Supplement 1).

Criterion validity

In general, weak correlations were observed between the IOI-AI^{op} factors and audiometry. A moderate negative correlation between postoperative PTA_4 AC and Factor 2^{op} (participation restrictions) (rho= −0.401, $p < 0.001$) was detected (Supplement 2).

Reliability

The reliability was assessed by calculating the internal consistency (Cronbach's alpha) and the intraclass correlation coefficient (ICC). The results indicated very strong agreement. The internal consistency was 0.83 and 0.82 for Factors 1^{op} and 2^{op}, respectively. The ICCs were 0.95 and 0.94 for Factors 1^{op} and 2^{op}, respectively (Table 5).

Discussion

The aim of the present study was to analyse the psychometric properties of the Swedish IOI-AI adapted to ear surgery (IOI-AI^{op}). The questionnaire was tested on individuals with otosclerosis after the stapedotomy intervention. In general, high mean

Table 5. Reliability of the IOI-AI^{op}.

IOI-AI^{op} subscales	Test (n = 21) (Mean ± SD)	Retest (n = 21) (Mean ± SD)	ICC (95% CI)[a]	Internal consistency reliability (95% CI)[b]
Factor 1^{op}	11.3 ± 3.3	11.5 ± 3.3	0.95 (0.89, 0.98)	0.83 (0.78, 0.87)
Factor 2^{op}	12.3 ± 2.8	12.2 ± 2.6	0.94 (0.87, 0.98)	0.82 (0.76, 0.86)

Groups: (STp) Stapedotomy, without prior hearing aid use or prior ear surgery – primary intervention ($n = 47$); (STs) Stapedotomy, with prior hearing aid use – secondary intervention ($n = 29$); (STq) Stapedotomy, included from the Swedish quality register for otosclerosis surgery – quality register ($n = 86$).
[a]Test-retest reliability measured by intraclass correlation coefficient ($n = 21$).
[b]Internal consistency measured by Cronbach's alpha.

scores were noted for all items, indicating a favourable outcome for the stapedotomy performed. Differences were encountered between the study groups, with the most favourable outcome for group STs with bilateral hearing loss and prior hearing aid use. The differences were focused on items included in Factor 1^{op} (benefit, satisfaction and quality of life) while items included in Factor 2^{op} (RAL, RPR and impact on others) were similar across the different groups. In the IOI-HA, Factor 1 was interpreted by Stephens in 2002 as representing satisfaction, while Factor 2 represents residual problems (Stephens 2002). There are probably several factors affecting the results of our study, including differences in uni- or bilateral hearing loss, the study design with one cohort being prospectively included at the clinic prior to surgery and one cohort cross sectionally included from the Quality Register for Otosclerosis Surgery 1 year after stapedotomy.

The frequency distribution of ratings showed the highest scores for the most favourable outcome (Robinson, Gatehouse, and Browning 1996) in all items except for the RAL and quality of life items (Figure 1). The frequency distribution has similarities to the pattern demonstrated by Brännström & Wennerström in a Swedish hearing-aid outcome study using the IOI-HA as an outcome measure (Brännström and Wennerström 2010). Hickson et al. evaluated a communication program using IOI-AI as an outcome measure. In the study by Hickson et al. the most favourable outcome was present for satisfaction and impact on others while all other items had the highest percentage of response in the mid-range (3–4 out of 5)(Hickson, Worrall, and Scarinci 2006). In the present study, the frequency pattern had a ceiling effect that was most pronounced for the satisfaction item with a 72.8% rating for the highest score. In a previous study from the Swedish quality registry data on otosclerosis surgery, 92.9% of the subjects reported better or much better hearing after surgery (Strömbäck et al. 2017). Stapedotomy is considered a relatively safe procedure with few complications and good hearing outcomes (Pauli et al. 2019). The good hearing outcomes could possibly explain the ceiling effect. However, the ceiling effect is a negative quality that can affect a questionnaire's ability to distinguish between different groups and treatments.

The type of hearing loss could also be a factor affecting scores on the questionnaire. All participants had conductive or mixed hearing loss prior to surgery, whereas in other studies, sensorineural hearing loss was the predominant type of hearing loss. Few studies have focused on conductive or mixed hearing loss. In a retrospective study assessing hearing aid use and benefit in a cohort of otosclerosis subjects 30 years after surgery, the IOI-HA scores were comparable to those obtained in this study (Redfors, Hellgren, and Möller 2013). This is consistent with the study by Brännström et al., where individuals with conductive or mixed hearing loss reported more favourable outcomes than the group with pure sensorineural hearing loss (Brännström and Wennerström 2010). However, in these studies, the IOI-HA was used and not the IOI-AI, so this comparison has to be made with caution.

The PCA was consistent with the original psychometric studies of IOI-HA and resulted in a two-component solution (Cox and Alexander 2002; Kramer et al. 2002). However, one cross-loading was identified. Item residual activity limitation (RAL) loaded equally on both Factor 1^{op} (0.62) and 2^{op} (0.61). Cross loadings for item RAL have been described in earlier studies (Heuermann, Kinkel, and Tchorz 2005; Stephens 2002). In the study by Heurmann, RAL loaded to Factor 1 in the mailing campaign group and to Factor 2 in the field test group. A hypothesis was that the question could be interpreted in different ways depending on when and how the question was asked (Heuermann, Kinkel, and Tchorz 2005). Another point of view was put forward by Manchaiah et al. debating whether cross loading could reflect the fact that each item represents a single construct in contrast to other questionnaires where several items form a construct (Manchaiah, Thammaiah, and Vinay 2021). It is possible that the study design in the present study, with different study cohorts, could have affected the cross loading for RAL. In our calculations we chose to add RAL to Factor 2^{op}, reflecting residual problems, as in the original studies by Cox & Alexander and Kramer (Cox and Alexander 2002; Kramer et al. 2002) as well as in the Swedish study by Brännström and Wennerström (2010). Criterion validity showed in general weak correlations. One exception was a moderate correlation between postoperative PTA^4 AC and Factor 2^{op} (residual difficulties). Postoperative hearing level (PTA^4 AC) reflects the actual hearing function and can be one of several factors affecting residual difficulties.

In contrast, a large Swedish quality register study by Arlinger et al., could not find any correlations between the IOI-HA total score and hearing level (measured as PTA). It was concluded that PTA is not a reliable measure of the benefit and satisfaction of hearing aid acquisition (Arlinger, Nordqvist, and Öberg 2017).

Convergent and discriminant validity was tested by correlating Factor 1^{op} and 2^{op} to other questionnaires, including the generic (SF36), the ORL intervention specific (GBI) and the hearing intervention specific (GHABP). As hypothesised, the strongest correlations were observed for the hearing-specific subscales of benefit and satisfaction while no correlations were found for scales measuring concepts that IOI-AI^{op} does not. Overall, the quality indicator of >75% agreement was fulfilled indicating good convergent and discriminant validity.

The questionnaire had excellent reliability scores. For Factor 1^{op} and Factor 2^{op}, the ICCs were 0.95 and 0.94, and the Cronbach's alpha coefficients were 0.83 and 0.82, respectively. High reliability scores have also been demonstrated in studies regarding IOI-HA, IOI-CI and IOI-AI (Heggdal et al. 2021; Hickson, Worrall, and Scarinci 2006; Smith, Noe, and Alexander 2009), indicating that the questionnaire's outcomes are reproducible and consistent.

A limitation to the study concerns the study population consisting of two cohorts, one prospectively sampled and one cross-sectionally sampled cohort. The subjects were invited to participate in different ways. The first cohort was recruited by the treating physician or audiologist, and the second cohort was

recruited by letter and/or a telephone call. This probably affected the results in a systematic unwanted way. A more favourable outcome could be expected if you were invited to the study by your treating surgeon compared to someone unknown.

Another limitation of the study is the distribution of the questionnaires. In the original study and most of the published studies, the questionnaires were administered by mail in a paper version. In this study, the questionnaires were sent by e-mail, and the responses were collected from a data platform. This could have affected the outcome. However, the mode of administration were both self-administered and visual and previous studies comparing web based versus paper and pencil self-administered questionnaires found fewer differences compared to face to face administration (Braekman et al. 2020; Bowling 2005).

Further studies are needed to assess how different factors affect the outcome. The IOI-AI^{op} is a comprehensive questionnaire that could be an option to include in the quality register for otosclerosis surgery.

Conclusion

The IOI-AI^{op} showed good psychometric properties and fulfilled the quality criteria proposed by Terwee et al. (2007). One disadvantage is the presence of ceiling effects. The PCA yielded a two-component solution, consistent with previous factor analyses of the IOI-HA. The comprehensive IOI-AI^{op} is, in our opinion, a useful tool to evaluate patient perspectives after ear surgery, both in a clinical setting but also in future studies and as a validated questionnaire in a quality register such as the Swedish register for otosclerosis surgery. However, further studies are needed to assess different factors affecting outcomes.

Acknowledgements

We wish to express our gratitude to statistician Helena Johansson, PhD for her statistical work. We also wish to thank the study coordinators, Pia Hallqvist and Carina Åberg.

Ethical approval

This study was approved by the Regional Ethical Review Board in Gothenburg, Sweden (Dnr 351-11, T921-14, T245-17, 2020-02526). All included subjects signed a written consent form before entering the study.

Informed consent

The included subjects signed written informed consent forms before entering the study.

Disclosure statement

The authors have no conflicts of interest to declare.

Funding

This study was financially supported by the Research and Development Council of Region Västra Götaland and the Gothenburg Medical Society.

ORCID

Ylva Dahlin Redfors http://orcid.org/0000-0002-1234-6573

Data availability statement

Data are available from the corresponding author upon reasonable request.

References

Arlinger, S. 2003. "Negative Consequences of Uncorrected Hearing Loss–A Review." *International Journal of Audiology* 42 (sup2): 17–20. https://doi.org/10.3109/14992020309074639.

Arlinger, S., P. Nordqvist, and M. Öberg. 2017. "International Outcome Inventory for Hearing Aids: Data from a Large Swedish Quality Register Database." *American Journal of Audiology* 26 (3S): 443–450. https://doi.org/10.1044/2017_AJA-16-0123.

Bowling, A. 2005. "Mode of Questionnaire Administration can have Serious Effects on Data Quality." *Journal of Public Health (Oxford, England)* 27 (3): 281–291. https://doi.org/10.1093/pubmed/fdi031.

Braekman, E., R. Charafeddine, S. Demarest, S. Drieskens, F. Berete, L. Gisle, J. Van der Heyden, and G. Van Hal. 2020. "Comparing Web-Based Versus Face-to-Face and Paper-and-Pencil Questionnaire Data Collected Through Two Belgian Health Surveys." *International Journal of Public Health* 65 (1): 5–16. https://doi.org/10.1007/s00038-019-01327-9.

Brännström, K. J., and I. Wennerström. 2010. "Hearing Aid Fitting Outcome: Clinical Application and Psychometric Properties of a Swedish Translation of the International Outcome Inventory for Hearing Aids (IOI-HA)." *Journal of the American Academy of Audiology* 21 (08): 512–521. https://doi.org/10.3766/jaaa.21.8.3.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*, 567. 2nd ed. Hillsdale, N.J.: L. Erlbaum Associates.

Cox, R. M., and G. C. Alexander. 1995. "The Abbreviated Profile of Hearing Aid Benefit." *Ear and Hearing* 16 (2): 176–186. https://doi.org/10.1097/00003446-199504000-00005.

Cox, R. M., and G. C. Alexander. 2002. "The International Outcome Inventory for Hearing Aids (IOI-HA): Psychometric Properties of the English Version." *International Journal of Audiology* 41 (1): 30–35. https://doi.org/10.3109/14992020209101309.

Cox, R. M., G. C. Alexander, and C. M. Beyer. 2003. "Norms for the International Outcome Inventory for Hearing Aids." *Journal of the American Academy of Audiology* 14 (08): 403–413. https://doi.org/10.1055/s-0040-1715761.

Dahlin Redfors, Y., R. Jönsson, and C. Finizia. 2022. "A Validation Study of the Swedish Version of the Glasgow Hearing Aid Benefit Profile Evaluated in Otosclerosis Subjects." *Laryngoscope Investigative Otolaryngology* 7 (3): 807–815. https://doi.org/10.1002/lio2.787.

Fayers, P. M. D. 2007. *Quality of Life: The Assessment, Analysis and Interpretation of Patient-Reported Outcomes*. 2nd ed. West Sussex, England: John Wiley & Sons.

Gandek, B., S. J. Sinclair, M. Kosinski, and J. E. Ware. Jr. 2004. "Psychometric Evaluation of the SF-36 Health Survey in Medicare Managed Care." *Health Care Financing Review* 25 (4): 5–25.

Gatehouse, S. 1999. "Glasgow Hearing Aid Benefit Profile: Derivation and Validation of a Client-centered Outcome Measure for Hearing Aid Services." *Journal of the American Academy of Audiology* 10 (02): 80–103. https://doi.org/10.1055/s-0042-1748460.

Gatehouse, S., and W. Noble. 2004. "The Speech, Spatial and Qualities of Hearing Scale (SSQ)." *International Journal of Audiology* 43 (2): 85–99. https://doi.org/10.1080/14992020400050014.

Granberg, S., J. Dahlström, C. Möller, K. Kähäri, and B. Danermark. 2014. "The ICF Core Sets for Hearing Loss–Researcher Perspective. Part I: Systematic Review of Outcome Measures Identified in Audiological Research." *International Journal of Audiology* 53 (2): 65–76. https://doi.org/10.3109/14992027.2013.851799.

Heggdal, P. O. L., M. H. Naess, J. Hess-Erga, K. S. Larsen, and H. J. Aarstad. 2021. "Psychometric Properties for the Norwegian Translations of Two Revised APHAB-Subscales and an Adapted IOI-HA (IOI-CI) in Patients with Cochlear Implants." *International Journal of Audiology* 61: 1–8.

Heuermann, H., M. Kinkel, and J. Tchorz. 2005. "Comparison of Psychometric Properties of the International Outcome Inventory for Hearing Aids (IOI-hA) in Various Studies." *International Journal of Audiology* 44 (2): 102–109. https://doi.org/10.1080/14992020500031223.

Hickson, L., L. Worrall, and N. Scarinci. 2006. "Measuring Outcomes of a Communication Program for Older People with Hearing Impairment Using the International Outcome Inventory." *International Journal of Audiology* 45 (4): 238–246. https://doi.org/10.1080/14992020500429625.

Hill-Feltham, P. R., M. L. Johansson, W. E. Hodgetts, A. V. Ostevik, B. J. McKinnon, P. Monksfield, R. Sockalingam, T. Wright, and J. R. Tysome. 2021. "Hearing Outcome Measures for Conductive and Mixed Hearing Loss Treatment in Adults: A Scoping Review." *International Journal of Audiology* 60 (4): 239–245. https://doi.org/10.1080/14992027.2020.1820087.

Kramer, S. E., G. H. Allessie, A. W. Dondorp, A. A. Zekveld, and T. S. Kapteyn. 2005. "A Home Education Program for Older Adults with Hearing Impairment and their Significant Others: A Randomized Trial Evaluating Short- and Long-Term Effects." *International Journal of Audiology* 44 (5): 255–264. https://doi.org/10.1080/14992020500060453.

Kramer, S. E., S. T. Goverts, W. A. Dreschler, M. Boymans, and J. M. Festen. 2002. "International Outcome Inventory for Hearing Aids (IOI-HA): Results from The Netherlands." *International Journal of Audiology* 41 (1): 36–41. https://doi.org/10.3109/14992020209101310.

Lisspers, J., A. Nygren, and E. Söderman. 1997. "Hospital Anxiety and Depression Scale (HAD): Some Psychometric Data for a Swedish Sample." *Acta Psychiatrica Scandinavica* 96 (4): 281–286. https://doi.org/10.1111/j.1600-0447.1997.tb10164.x.

Manchaiah V., S. Thammaiah, and Vinay. 2021. Psychometric Properties of the Kannada Version of the International Outcome Inventory for Hearing Aids (IOI-HA). *International Journal of Audiology* 60(12): 1039–1045. https://doi.org/10.1080/14992027.2021.1884910.

Noble, W. 2002. "Extending the IOI to Significant Others and to Non-Hearing-Aid-Based Interventions." *International Journal of Audiology* 41 (1): 27–29. https://doi.org/10.3109/14992020209101308.

Öberg, M., T. Lunner, and O. Andersson. 2007. "Psychometric Evaluation of Hearing Specific Self-Report Measures and Their Associations with Psychosocial and Demographic Variables." *Audiological Medicine* 5 (3): 188–199. https://doi.org/10.1080/16513860701560214.

Ostevik, A. V., P. Hill-Feltham, M. L. Johansson, B. J. McKinnon, P. Monksfield, R. Sockalingam, J. R. Tysome, T. Wright, and W. E. Hodgetts. 2021. "Psychosocial Outcome Measures for Conductive and Mixed Hearing Loss Treatment: An Overview of the Relevant Literature." *International Journal of Audiology* 60 (9): 641–649. https://doi.org/10.1080/14992027.2021.1872805.

Pauli, N., K. Strömbäck, L. Lundman, and Y. Dahlin-Redfors. 2019. "Surgical Technique in Stapedotomy Hearing Outcome and Complications." *Laryngoscope* 130 (3): 790–796. https://doi.org/10.1002/lary.28072.

Redfors, Y. D., J. Hellgren, and C. Möller. 2013. "Hearing-Aid Use and Benefit: A Long-Term Follow-Up in Patients Undergoing Surgery for Otosclerosis." *International Journal of Audiology* 52 (3): 194–199. https://doi.org/10.3109/14992027.2012.754957.

Redfors, Y. D., R. Jönsson, B. Tideholm, and C. Finizia. 2019. "Psychometric Properties of the Swedish Version of the Glasgow Benefit Inventory in Otosclerosis Subjects." *Laryngoscope Investigative Otolaryngology* 4 (6): 673–677. https://doi.org/10.1002/lio2.320.

Robinson, K., S. Gatehouse, and G. G. Browning. 1996. "Measuring Patient Benefit from Otorhinolaryngological Surgery and Therapy." *Annals of Otology, Rhinology, and Laryngology* 105 (6): 415–422. https://doi.org/10.1177/000348949610500601.

Smith, S. L., C. M. Noe, and G. C. Alexander. 2009. "Evaluation of the International Outcome Inventory for Hearing Aids in a Veteran Sample." *Journal of the American Academy of Audiology* 20 (6): 374–380. https://doi.org/10.3766/jaaa.20.6.5.

Stephens, D. 2002. "The International Outcome Inventory for Hearing Aids (IOI-HA) and its Relationship to the Client-Oriented Scale of Improvement (COSI)." *International Journal of Audiology* 41 (1): 42–47. https://doi.org/10.3109/14992020209101311.

Stewart, A. W. J. 1992. *Measuring Functioning and Well-Being: The Medical Outcomes Study Approach*. Durham, North Carolina: Duke University Press.

Strömbäck, K., L. Lundman, A. Bjorsne, J. Grendin, A. Stjernquist-Desatnik, and Y. Dahlin-Redfors. 2017. "Stapes Surgery in Sweden: Evaluation of a National-Based Register." *European Archives of Oto-Rhino-Laryngology: official Journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS): Affiliated with the German Society for Oto-Rhino-Laryngology – Head and Neck Surgery* 274 (6): 2421–2427. https://doi.org/10.1007/s00405-017-4510-2.

Sullivan, M., and J. Karlsson. 1998. "The Swedish SF-36 Health Survey III. Evaluation of Criterion-Based Validity: Results from Normative Population." *Journal of Clinical Epidemiology* 51 (11): 1105–1113. https://doi.org/10.1016/s0895-4356(98)00102-4.

Sullivan, M., J. Karlsson, and J. E. Ware. Jr. 1995. "The Swedish SF-36 Health Survey–I. Evaluation of Data Quality, Scaling Assumptions, Reliability and Construct Validity Across General Populations in Sweden." *Social Science & Medicine (1982)* 41 (10): 1349–1358. https://doi.org/10.1016/0277-9536(95)00125-q.

Taft, C., J. Karlsson, and M. Sullivan. 2004. "Performance of the Swedish SF-36 Version 2.0." *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation* 13 (1): 251–256. https://doi.org/10.1023/B:QURE.0000015290.76254.a5.

Terwee, C. B., S. D. M. Bot, M. R. de Boer, D. A. W. M. van der Windt, D. L. Knol, J. Dekker, L. M. Bouter, and H. C. W. de Vet. 2007. "Quality Criteria were Proposed for Measurement Properties of Health Status Questionnaires." *Journal of Clinical Epidemiology* 60 (1): 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012.

Wu, Y., B. Levis, Y. Sun, C. He, A. Krishnan, D. Neupane, P. M. Bhandari, Z. Negeri, A. Benedetti, and B. D. Thombs, DEPRESsion Screening Data (DEPRESSD) HADS Group. 2021. "Accuracy of the Hospital Anxiety and Depression Scale Depression Subscale (HADS-D) to Screen for Major Depression: Systematic Review and Individual Participant Data Meta-Analysis." *BMJ (Clinical Research ed.)* 373:n972. https://doi.org/10.1136/bmj.n972.

Zahnert, T., H. Löwenheim, D. Beutner, R. Hagen, A. Ernst, H.-W. Pau, T. Zehlicke, H. Kühne, N. Friese, A. Tropitzsch, et al. 2016. "Multicenter Clinical Trial of Vibroplasty Couplers to Treat Mixed/Conductive Hearing Loss: First Results." *Audiology & Neuro-Otology* 21 (4): 212–222. https://doi.org/10.1159/000444616.

Zigmond, A. S., and R. P. Snaith. 1983. "The Hospital Anxiety and Depression Scale." *Acta Psychiatrica Scandinavica* 67 (6): 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x.