Accepted: 29 October 2017

DOI: 10.1111/coa.13026

ORIGINAL ARTICLE

WILEY

Quality of questionnaires for the assessment of otitis media with effusion in children

R.W.C. Gan[1] | M. Daniel[2] | M. Ridley[3] | J.G. Barry[1,3]

[1]Nottingham University Hospitals NHS Trust, Queen's Medical Centre, Nottingham, UK

[2]ENT Department, Nottingham University Hospitals, Nottingham, UK

[3]Medical Research Council Institute of Hearing Research, School of Medicine, The University of Nottingham, Nottingham, UK

Correspondence
R.W.C. Gan, Nottingham University Hospitals NHS Trust, Queen's Medical Centre, Nottingham, UK.
Email: richardwcgan@gmail.com

Funding information
The University of Nottingham Medical Research Council, Grant/Award Number: U135097130

Objectives: Audiometric tests provide information about hearing in otitis media with effusion (OME). Questionnaires can supplement this information by supporting clinical history-taking as well as potentially providing a standardized and comprehensive assessment of the impact of the disease on a child. There are many possible candidate questionnaires. This study aimed to assess the quality and usability of parent / child questionnaires in OME assessment.

Design and main outcome measures: Fifteen, published questionnaires, commonly used in audiological departments (Auditory Behaviour in Everyday Life (ABEL), Children's Auditory Performance Scale (CHAPS), Children's Home Inventory for Listening Difficulties (CHILD), Children's Outcome Worksheets (COW), Evaluation of Children's Listening and Processing Skills (ECLiPS), Early Listening Function (ELF), Fisher's Auditory Problem Checklist (FAPC), Hearing Loss 7 (HL-7), Listening Inventory for Education- Revised (LIFE-R Student), Listening Inventory for Education UK Individual Hearing Profile (LIFE-UK IHP), LittlEARS Auditory Questionnaire (LittlEARS), Listening Situations Questionnaire (LSQ), Otitis Media 6 (OM-6), Quality of Life in Children's Ear Problems (OMQ-14), Parents' Evaluation of Aural/Oral Performance of Children (PEACH) were assessed according to the following 8 criteria: conceptual clarity, respondent burden, reliability, validity, normative data, item bias, ceiling/ floor effects, and administrative burden.

Results: ECLiPS, LittlEARS and PEACH scored highest overall based on the assessment criteria established for this study. None of the questionnaires fully satisfied all 8 criteria. Although all questionnaires assessed issues considered to be of at least adequate relevance to OME, the majority had weaknesses with respect to the assessment of psychometric properties, such as item bias, floor/ceiling effects or measurement reliability and validity. Publications reporting on the evaluation of reliability, validity, normative data, item bias and ceiling/floor effects were not available for most of the questionnaires.

Conclusions: This formal evaluation of questionnaires, currently available to clinicians, highlights three questionnaires as potentially offering a useful adjunct in the assessment of OME in clinical or research settings. These were the ECLiPS, which is suitable for children aged 6 years and older, and either the LittlEARS or the PEACH for younger children. The latter two are narrowly focused on hearing, whereas ECLiPS has a broader focus on listening, language and social difficulties.

1 | INTRODUCTION

Otitis media with effusion (OME) is the commonest cause of hearing loss among children. Studies on British children have found prevalence as high as 36.6% among 8-month-old children,[1] an annual prevalence of 16% among 5-year-olds,[2] which decreases to 6% in those aged 8 years.[2] These studies have also shown that OME is more common in the winter months.[1] Although the natural history is towards self-resolution,[2] in some cases, it persists beyond a few months and may require active treatment to prevent language, learning and behavioural problems. Treatment options include hearing aids, ventilation tube insertion or autoinflation.[3,4]

Decisions about treatment in the United Kingdom are supported by NICE guidelines which recommend that ventilation tube insertion should be considered for OME persisting at least 3 months and with hearing loss of 25 to 30 dB HL or worse in the better ear, or if there is a significant impact on the child's social, educational or developmental outcome.[5] Thus, in addition to objective tests such as the pure-tone audiogram, clinical history through parental report is used in decision-making, and questionnaires could offer a formalised means for obtaining this information. Decision-making in OME often does not follow guidelines,[6] and published studies often report a variety of different outcomes[7]; hence, identification of high-quality tools to capture parental report of the impact of OME is crucial to facilitate good and transparent decision-making both in the clinic and research setting.

In addition to supporting clinical history-taking, questionnaires can be useful tools, as they provide a standardised and comprehensive method of assessing the impact of the hearing loss on the everyday functioning of the child over a period of time. As the incidence and impact of OME fluctuate throughout the year,[8] such report-based measures provide additional benefit by capturing long-term effects.

Many different questionnaires have been developed for assessing children with hearing loss and associated difficulties. Some like the Otitis Media 6 (OM-6)[9] or the Quality of Life in Children's Ear Problems (OMQ-14)[10] have been specifically designed to support clinical history-taking for children with OME. Others have been developed for hearing screening purposes, or to target problems specific to the auditory periphery or other problems associated with the processing of auditory inputs (ie [central] auditory processing disorder [APD], sometimes referred to as listening difficulties), but as children with hearing loss share symptoms with children with OME, they may still be useful in the context of OME assessment.

The aim of this study was to systematically review questionnaires that are commonly used in the clinic to assess problems with either hearing- or listening-related difficulties in the context of clinical assessment of otitis media with effusion.

Keypoints

- There are many questionnaires available to aid assessment of otitis media with effusion (OME) in children.
- There are considerable gaps in the formal evaluation of available questionnaires.
- Evaluation of Children's Listening and Processing Skills (ECLiPS), LittlEARS Auditory Questionnaire (LittlEARS), and Parents' Evaluation of Aural/Oral Performance in Children (PEACH) satisfy most of the criteria laid out in this study for the evaluation of questionnaires.
- Although they can be useful adjuncts to audiometry, further research is needed to determine how they fit into the overall assessment of children with OME.

2 | METHODS

This study focused on assessing publicly-available questionnaires that collected carer or child (self) report-based assessment of hearing or listening abilities in children aged below 18 years. These questionnaires were identified through a Pubmed search using the keywords [listening difficulty] OR [hearing loss] OR [hearing impairment]; AND [survey] OR [questionnaire]; AND [child]. All the questionnaires that were identified were included, and a flow chart detailing excluded studies is therefore not provided.

Quality of questionnaires was assessed using a modification of the criteria proposed by Andresen who outlined a systematic process for critically assessing the properties of a questionnaire from the perspective of both measurement properties and usability.[11] This framework of criteria was subsequently applied by Bagatto et al.[12] as part of the process of developing guidelines for evaluating the outcome of hearing aid use in children. Bagatto et al's work established the usability of Andresen's framework in the hearing context. We further adapted this framework as described below to achieve a similarly structured assessment of questionnaires in the context of OME.

Bagatto et al applied 13 criteria including conceptual clarity (captures relevant domains ie symptoms of disease together with impacts on social, educational and developmental outcomes, normative data, measurement model (ceiling or floor effects), item or instrument bias, respondent burden (number of items, interpretability of items and usability of response scale), administrative burden (ease of scoring and interpretation), reliability (consistency of results across time and scorers), discriminant validity, convergent validity, responsiveness (ie sensitivity to treatment-related change), alternate or accessible forms (eg paper vs electronic), and adaptations for use in different languages or cultures). While it would have been interesting to assess treatment-related responsiveness, even the questionnaires specifically designed to look at treatment effects (see Children's Outcome Worksheets (COW) and Early listening Function (ELF), Table 1) did not formally assess this, and responsiveness was therefore excluded. We also did not assess whether there were alternative formats or other adaptations as there was little or no information available for any of the questionnaires assessed here.

Once the final framework of eight relevant categories was established, the questionnaires were reviewed and scored with respect to these categories (Tables 1 and 2). The ratings of some characteristics involved a subjective assessment on the part of the authors, while

TABLE 1 Summary of the fifteen questionnaires included in the study

Questionnaire	Purpose and domains	Target age group	Administration	Length (number of items)	Scoring	Reliability and validity	Normative data
Auditory Behaviour in Everyday Life (ABEL)	Assesses parent's perceptions of child's auditory behaviour. Domains covered include auditory awareness and verbal communication.	2-12 years	Parent indicates frequency that they observe behaviours specified in each item.	24 items	7-point ordinal scale: "never," "hardly ever," "occasionally," "about half the time," "frequently," "almost always" and "always"	Cronbach's alpha 0.94. Factor loading equal or greater than 0.49.	28 parents of children aged 4 to 14 years old (Mean=8; *Standard Deviation*=3) who had varying degrees of hearing loss ranging from mild, moderate to profound. Five implant users and 23 hearing aid users.
Children's Auditory Performance Scale (CHAPS)	Assesses listening in a variety of conditions. Domains covered include listening in noise, quiet, ideal, multiple inputs, as well as auditory memory sequencing, and auditory attention span.	7 years+	Parent (or teacher) indicates level of listening difficulty compared to children of the same age and background. There is additional space for open comments at the end of each section.	36 items	7-point scale: less difficulty (+1), same amount (0), slightly more (−1), considerably more (−2), significantly more (−4) and can't function (−5)	Not reported	40 children (20 with (central) auditory processing disorder ([C]APD), 20 without (C)APD, of similar age and background.
Children's Home Inventory for Listening Difficulties (CHILD)	Assesses hearing and understanding in a variety of conditions.	3-12 years	Parents (or other family members) rate hearing and understanding for each situation. Children aged 7 or older can also do this.	15 items (situations)	8-point scale: "Huh?," "tough going," "sometimes get it, sometimes don't," "it takes work but usually can get it," "okay but not easy," "pretty good," "good" and "great."	Not reported	Not reported
Children's Outcome Worksheets (COW)	Assesses impact of hearing loss and improvement subsequent to provision of amplification. Hearing functionality is assessed by the parent child and teacher for a minimum of three situations, (or possibly more if nominated by the respondent).	4-12 years	Multiple observers possible, including parent, teacher and child (self). Each situation is described for the listener-specific circumstances then rated for (1) importance, (2) (degree of change (post-amplification) (3) ability (with hearing instrument).	Minimum of 3 situations per observer (teacher, parent, child), with additional writing space for other situations of relevance.	5-point scale × 2: Degree of change – "worse (1)," "no difference (2)," "slightly better (3)," "better (4)," "much better (5)" Ability (with hearing instrument) – "hardly ever 10% (1)," "occasionally 25% (2)," "half the time 50% (3)," "Most of time 75% (4)," "Almost always 95% (5)"	Not reported	Not reported

(Continues)

TABLE 1 (Continued)

Questionnaire	Purpose and domains	Target age group	Administration	Length (number of items)	Scoring	Reliability and validity	Normative data
Evaluation of Children's Listening and Processing Skills (ECLiPS)	Assesses severity of symptoms associated with listening difficulty in children referred for suspected auditory processing disorder. Domains covered include Speech & Auditory Processing, Environmental & Auditory Sensitivity, Language/Literature/ Laterality, Pragmatic & Social Skills, and Memory & Attention.	6-11 years	Parents rate statements about symptoms	38 items	5-point (thermometer colours) Likert scale: "Strongly Agree," "Agree," "Neither Agree nor disagree," "Disagree," "Strongly Disagree"	Cronbach's alpha 0.83-0.94; Test-retest reliability 0.9-0.96; Inter-rater reliability (parent-parent) 0.78-0.88/ (parent-teacher) 0.4-0.7; Construct validity, assessed by comparison with CHAPS (hearing), CCC-2 (language), SCQ (autism). Performed as predicted with the related and unrelated subscales on the CHAPS, CCC-2, and SCQ.	660 typically-developing children aged 6 to 11 years old.
Early listening Function (ELF)	Assesses (a) hearing functionality in very young children with hearing loss and (b) impact of amplification on hearing. Hearing is assessed for three kinds of activities: quiet, typical and loud.	5 months – 3 years	Completed by parent/ caregiver. Involves administration of auditory stimuli with the child placed at different distances from the sound source. The parent/caregiver rates responsiveness for each location.	9 items 12 listening activities.	3-point scale: "Yes," "No," "Maybe."	Not reported	Not reported
Fisher's Auditory Problem Checklist (FAPC)	Assesses severity of symptoms associated with listening difficulty in children referred for suspected auditory processing disorder. Domains include attention, listening, speech, language, memory, academic difficulty.	Not clear but presumably for children in formal education only.	Parent (or teacher) indicates which symptoms from the list are observed in the child.	25 items	Checklist – space for a mark next to items (symptoms) relevant to the child.	Not reported	Not reported

(Continues)

TABLE 1 (Continued)

Questionnaire	Purpose and domains	Target age group	Administration	Length (number of items)	Scoring	Reliability and validity	Normative data
Hearing Loss 7 (HL-7)	Assesses children for symptoms of hearing loss. Domains include behaviours associated with symptoms associated with hearing loss.	6 months - 12 years	Parents indicate frequency that they observe each symptom.	7 items	7-point ordinal scale: "none of the time," "hardly any of the time," a little of the time," "some of the time," "a good bit of the time," "most of the time" and "all of the time"	Test-retest reliability 0.7; Internal consistency (item-total correlation) "fair to good"/"high internal consistency"; Construct validity 0.67 ($P < .001$); Criterion validity 0.10 ($P = .29$).	115 parents of predominantly male children (64%), aged between 0.9 and 12.6 years with OME.
Listening Inventory For Education – Revised (LIFE-R Student)	Identifies classroom situations, which present listening challenges. Domains: Hearing, understanding.	8 years+	Students rate their hearing and understanding in different classroom situations.	15 items	5-point scale: "always easy," "mostly easy," "sometimes difficult," "mostly difficult," "always difficult."	Not reported	Not reported
Listening Inventory for Education UK Individual Hearing Profile (LIFE-UK IHP)	To identify listening challenges in the UK classroom setting. Domains: Hearing	7-14 years	Child rates hearing ability in various situations.	18 items with visual cues to clarify meaning	5-point scale with visual cues to meaning: "always easy," "mostly easy," "sometimes difficult," "mostly difficult," "always difficult."	Not reported	Not reported
LittlEARS Auditory Questionnaire (LittlEARS)	Assesses auditory behaviour in the pre-verbal stage of a child's life. Domains include auditory responsiveness to stimuli like music, speech, environmental sounds, as well as, capacity to connect sound to source.	0-2 years	Independent parent/ caregiver (or supported during clinician interview) respond to series of closed questions.	35 items	"Yes," "No"	Cronbach's alpha 0.96	218 normal hearing children aged between 5 days and 24 months old.
Listening Situations Questionnaire (LSQ)	Assesses listening in various situations. Domain: Not specified. Understood to be hearing/understanding.	6 years+	Parents rate different listening situations according to: (a) importance, (b) frequency, (c) difficulty. Hearing aid use is specified for this last part, but we treated as general difficulty.	10 situations, 3 rating scales per situation plus 1 free text question	Two 4-point rating scales varying with item: for example, "very important" to "not at all important" or "very often" to "hardly ever." One 5-point scale: "lots and lots" to "doesn't wear hearing aid"	Not reported	Not reported

(Continues)

TABLE 1 (Continued)

Questionnaire	Purpose and domains	Target age group	Administration	Length (number of items)	Scoring	Reliability and validity	Normative data
Otitis Media–6 (OM-6)	Assesses severity of symptoms associated with OME. The 6 domains assessed are Physical suffering, Hearing loss, Speech impairment, Emotional distress, Activity limitations and Caregiver concerns.	Young children. Age range not specified.	Parent/caregiver indicates how much each domain has been impacted by the disease in the preceding 3 month.	6 items	7-point categorical scale: "Not present/no problem," "Hardly a problem at all," "Somewhat of a problem," "Moderate problem," "Quite a bit of a problem," "Very much of a problem," "Extreme problem."	Cronbach's alpha 0.75. Good construct validity based on 87.5% of hypothesised correlations between items (inter-item), between items and summary scores (item-total) and between summary scores (total-total) within and between the questionnaires.	Not known
Quality of Life in children's ear problems (OMQ-14)	Assesses the impact of OME on quality of life and auditory outcomes. Domains include Physical health, ear problems, hearing, attention, speech, parental well-being.	3-9 years	Parent/caregiver respond using a pre-defined response options, to a series of open or closed questions about symptoms observed in the preceding 3 months.	14 items	Response options vary with the item. For example, "very good," "good," "only fair" or "poor" "Not at all," "once, 2-3 times," "4 or more times"	Not reported.	Not reported.
Parents' Evaluation of Aural/Oral Performance of Children (PEACH)	Assesses hearing and listening post provision of hearing prosthesis (eg hearing aid; cochlear implant). Domains include ability to hear and communicate (in quiet, in different noisy situations, and when using a telephone), and responsiveness to different sounds in the environment.	Authors do not state an age limit. Items suited to preschool up to 7 years	Parents rate frequency of behaviours specified in the questionnaire.	13 situations/ conditions	5-point scale: "Never," "Seldom," Sometimes', "Often," "Always"	Cronbach's alpha 0.88.	90 parents of children with normal hearing and 90 parents of children with hearing impairment

APD, auditory processing disorder.

TABLE 2 Summary of questionnaire ratings for each of the assessment criteria

Questionnaire	Conceptual clarity	Respondent burden	Reliability	Validity	Normative data	Item bias	Ceiling/floor effects	Administrative burden
ABEL	B	A	A	B	C	U	U	B
CHAPS	B	C	U	U	U	U	U	B
CHILD	B	B	U	U	U	U	U	A
COW	C	B	U	U	U	U	U	B
ECLIPS	B	A	A	B	B	B	A	C
ELF	B	A	U	U	U	U	U	B
FAPC	A	A	U	U	U	U	U	A
HL-7	A	A	B	U	C	U	U	A
LIFE-R Student	B	B	U	U	U	U	U	C
LIFE-UK IHP	B	B	U	U	U	U	U	C
LittlEARS	B	A	A	U	B	A	U	A
LSQ	C	B	U	U	B	U	C	C
OM-6	A	A	A	U	U	U	U	A
OMQ-14	A	A	U	U	U	U	U	A
PEACH	B	A	A	U	A	U	U	B

ABEL, Auditory Behaviour in Everyday Life; CHAPS, Children's Auditory Performance Scale; CHILD, Children's Home Inventory for Listening Difficulties; COW, Children's Outcome Worksheets; ECLiPS, Evaluation of Children's Listening and Processing Skills; ELF, Early Listening Function; FAPC, Fisher's Auditory Problem Checklist; HL-7, Hearing Loss 7; LIFE-R Student, Listening Inventory for Education- Revised; LIFE-UK IHP, Listening Inventory for Education UK Individual Hearing Profile; LittlEARS, LittlEARS Auditory Questionnaire; LSQ, Listening Situations Questionnaire; OM-6, Otitis Media 6; OMQ-14, Quality of Life in Children's Ear Problems; PEACH, Parents' Evaluation of Aural/Oral Performance of Children.

the ratings of other criteria involved reference to published materials about the questionnaires to determine whether specific objective measures had been collected.

All authors independently scored 3 randomly selected questionnaires first, with the final rating for each characteristic determined through a consensus discussion. This initial rating was also used to ensure agreement on the criteria of rating. The remaining questionnaires were rated independently by two of the authors with final rating for each characteristic determined through consensus discussion. For each characteristic (Table 2), a rating of excellent "A," moderate "B" or poor "C" was allocated. Where information was lacking, the code "U," meaning unavailable, was applied.

2.1 | Conceptual clarity

The questionnaire should be developed with the purpose of capturing information regarding the disease's impact not only in relation to the child's perceived level of hearing, but also their speech, language, psychosocial development[5,13] and capacity to function in educational settings.[5] Ratings were based on the authors' evaluation. A rating of excellent "A" was awarded if the questionnaire had a completely relevant purpose, moderate "B" if it captured some issues of relevance, but not all and poor, "C," if it was only tangentially relevant.

2.2 | Respondent burden

Reliability of response depends to some extent on respondent burden. On a simplistic level, burden can be defined according to numbers of items. However, of more significance is the ease with which items can be understood and rated by the respondent. Brevity and ease of understanding were, therefore, both considered. Two volunteer parents completed the questionnaires and rated each in terms of completion time, number of questions, clarity of instructions and items, as well as providing an overall assessment of how happy they would have been to complete the questionnaire as part of a clinical assessment. A questionnaire was rated "A" if both respondents largely agreed a questionnaire was acceptable in terms of respondent burden. It was rated "B" if there was disagreement regarding ease of understanding and completion. It was rated "C" if both volunteers agreed on there being significant problems with usability.

2.3 | Reliability

The questionnaires should provide consistent results. Reliability can be further split into test-retest reliability, inter-rater reliability and internal consistency (Cronbach's alpha). These three measures were available for the Evaluation of Children's Listening and Processing Skills (ECLiPS). Scores for the remaining questionnaires were based on whatever was reported (typically Cronbach's alpha or test-retest reliability). If the necessary data and the statistics assessing test-retest, inter-rater reliability and internal consistency were unavailable, the questionnaire was given a score of "U."

Test-retest reliability is how consistent the results are after multiple administrations to the same group of respondents. This is statistically measured using the retest intraclass correlation coefficient (r). Questionnaires were rated "A" if r was greater than or equal to .75, "B" if it was between .75 and .40, and "C" if it was less than or equal to .40.

Internal consistency refers to how consistently items within the same factor measure the same construct.[12] This can be assessed statistically using Cronbach's alpha, a correlation coefficient. Questionnaires with a Cronbach's alpha greater than or equal to .8 were rated "A," less than .8 but more than .7 "B" and less than or equal to .7 "C."

Inter-rater reliability refers to how two different people score the same child. If the intraclass correlation was greater than or equal to .75 questionnaires were rated "A," "B" if it was between .75 and .40, and "C" if it was less than or equal to .40.

Reliability can be considered a precursor to validity,[14] as a test must be reliable to be valid.[15]

2.4 | Validity

Validity, specifically construct validity, refers to how well the questionnaire measures what it is supposed to be measuring. This is assessed in terms of measurement model. The measurement model is established through exploratory factor analysis which is a statistical method for identifying a few underlying latent variables (often called factors or constructs) that exist between a larger number of correlated variables.

Criterion-related and discriminant validity consider the nature of the underlying latent variables being measured.

Discriminant validity assesses the extent to which questionnaires that are predicted to measure different constructs are unrelated (ie do not correlate). Correlation coefficients are predicted to be non-significant if the constructs measured are sufficiently unrelated. Discriminant validity measures were only available for the ECLiPS. Validity scores for the other questionnaires are therefore based on observed criterion validity measures only.

Criterion-related validity considers whether the questionnaire measures the construct of interest as well as an already well-established questionnaire. This involves determining correlations between the questionnaire of interest and a known gold standard measure of the construct to be probed. Often such gold standard measures do not exist making it hard to assess criterion validity. However, if criterion validity was assessed, a rating of "A" was given for a correlation coefficient of greater than or equal to .60, a rating of "B" was assigned if it was below .60 but greater than .30, and a rating of "C" if it was less than or equal to .30.

2.5 | Normative data

A child's symptoms of disease impact may change with age as their speech, language, cognitive and psychosocial skills develop. As a consequence, any non-physical symptoms of disease should be normed for these changes. Questionnaires with normative data for a sufficiently large number of children for reliable regression analyses with age (ie 20 – 30 children per age band), spread across different ages, with normal hearing and those with hearing loss were given a score "A"; "B" if the data were only available for children with normal hearing; "C" if only some preliminary experimental data were available; and "U" if no data were available.

2.6 | Item bias

Items within the questionnaire should not show bias towards a particular culture, race or gender. This can be assessed using a variety of techniques including simple regression analyses or more sophisticated modern psychometric techniques such as Rasch analysis. Questionnaires were rated "A" if there was evidence that bias effects were assessed, with sufficient numbers of participants to be able to do this reliably, and any effects observed were addressed as appropriate. Questionnaires were rated "B" if there was evidence of an attempt to assess bias effects, but there were insufficient numbers of participants to be able to do this reliably. They were rated "C" if there was acknowledgement that bias may affect observations, but no attempt was made to assess this, possibly because of small participant numbers. They were rated "U" if the issue of bias was not considered during development.

2.7 | Ceiling/floor effects

To be maximally sensitive to as many individual differences as possible, questionnaires should be minimally susceptible to ceiling or floor effects. This is evident in a marked tendency for response distributions to be skewed either to top or bottom of the response scale. Questionnaires were scored based on published evaluations. They were rated "A" if little skew to the extremes of the response scale was reported; "B" if there was some evidence for skewing; "C" if there was substantial skew to the top or bottom of the response scale, and "U" if the data describing response distributions were unavailable.

2.8 | Administrative burden

Administration of questionnaires should have a low administrative burden. That is results should be easily obtained, quantified and interpreted by the clinician. This was based on the subjective assessment of the authors, giving a score of "A" if the questionnaire involved simple addition and outputs were easy to interpret and interpret; "B" if scoring involved some additional manipulations but outputs were still easy to interpret; and "C" if scoring involved considerable manipulation of data and/or interpretation of outputs was not immediately obvious.

2.9 | Ethical considerations

Approval of an ethics committee was not required as analyses were based on published anonymised data.

3 | RESULTS

Fifteen questionnaires were identified and were assessed and rated (Tables 1 and 2). Ratings for each questionnaire according to the eight defined characteristics are summarised in Table 2.

3.1 | Conceptual clarity

Reflecting our search criteria, most questionnaires covered domains that were considered of at least adequate relevance to OME. OMQ-14 and OM-6 were specifically designed to assess children with OME and, as such, they captured most aspects of relevance to the disease and were rated "A."

The majority of the remaining questionnaires captured some aspects but not all of interest to clinicians assessing children with OME. They were therefore rated "B." Auditory Behaviour in Everyday Life (ABEL), Children's Home Inventory for Listening Difficulties (CHILD), HL-7, Parents' Evaluation of Aural/Oral Performance of Children (PEACH), LIFE-R Student and LIFE-UK IHP focused on capturing symptoms of hearing loss in everyday contexts. This is very useful for supplementing audiometric measurements in OME, as levels of hearing loss fluctuate. LittlEARS and ELF also assess hearing loss but have been specifically designed to probe for these symptoms in children younger than 2 years of age. Children's Auditory Performance Scale (CHAPS), ECLiPS and Fisher's Auditory Problem Checklist (FAPC) were purposed towards aiding the assessment of children with suspected (central) auditory processing disorder ([C] APD). In the context of OME, they provide a means for assessing the impact of OME on cognition and listening (ECLiPS; CHAPS; FAPC), speech/language (ECLIPS), psychosocial function (ECliPS) and functionality in educational settings (FAPC; ECLiPS).

Listening Situations Questionnaire and COW were rated "C" because they focused solely on hearing in a limited range of settings as well as being primarily designed to assess the benefit of amplification on hearing function.

3.2 | Respondent burden

The two volunteer parents found the ABEL, ECLiPS, ELF, FAPC, HL-7, LittlEARS, OM-6, OMQ14 and PEACH acceptable in terms of respondent burden. Both found the CHAPS difficult to use and understand. They had mixed views about the ease of use of the CHILD, COW, LIFE-R Student, LIFE-UK IHP and Listening Situations Questionnaire (LSQ).

Data for reliability, validity, normative data, item bias and ceiling/floor effects were unavailable for most questionnaires.

3.3 | Reliability

Parents' Evaluation of Aural/Oral Performance of Children and OM-6 had test-retest intraclass correlation coefficients of 0.93[16] and 0.85,[17] respectively. ECLiPS assesses 5 domains associated with listening, language and cognition and generates separate scores for each. Excellent test-retest reliability is reported (intraclass correlations range from .9 to .96[18]), as well as very good parent-parent inter-rater reliability (r between .78 and .88), but poorer parent-teacher inter-rater reliability (r between .4 and .7). Adequate test-retest reliability is reported for HL-7, with an intraclass correlation coefficient of .70.[19]

Excellent internal consistency is reported for PEACH (Cronbach's alpha .88),[16] ABEL (.94),[20] LittlEARS (.96),[21] ECLiPS (.83 to .94).[18] OM-6 has an adequate internal consistency with a Cronbach's alpha of .75.[17] This lower score may reflect the fact that Cronbach's alpha is influenced by the number of items included in the analysis, regardless of how closely the items relate to each other.

3.4 | Validity

Evaluation of Children's Listening and Processing Skills had good criterion-related and discriminant validity. Where ECLiPS domains were predicted to measure the same constructs as other questionnaires correlations, r of greater than .50[18] were observed (ie evidence of criterion validity). Likewise, where ECLiPS domains were predicted to measure different constructs to those measured by other questionnaires (ie evidence of discriminant validity), weak or no correlation was observed ($r < .35$).[18]

OM-6 was described as having good construct validity based on 87.5% of hypothesised correlations between items (inter-item), between items and summary scores (item-total) and between summary scores (total-total) within and between the questionnaires.[17] Factor loadings from factor analysis of ABEL also suggested good internal validity with most items having factor loadings greater than or equal to .49. Three items had factor loadings between .44 and .46.[20]

3.5 | Normative data

Parents' Evaluation of Aural/Oral Performance of Children had published normative data of a large number of normal hearing children and those with hearing/listening difficulties.[16] ECLiPS and LittlEARS had published normative data for normal hearing children.[22,23] ABEL reports data for a small number (28) of cochlear implant and hearing aid users.[20] LSQ looked at differences in responses on different items for older and younger children as well as normally-hearing and hearing-impaired children[24] though to our knowledge the data were not normed per se.

3.6 | Item bias

Evaluation of Children's Listening and Processing Skills and LittlEARS have published analysis, assessing of gender (ECLiPS, LittlEARS) socio-economic status (ECLiPS) on responses.[22,23,25] Scoring was adjusted to account for gender effects in the ECLiPS. No effects of gender were reported for LittlEARS.

3.7 | Ceiling/floor effects

Only ECLiPS reports data regarding floor or ceiling effects. This was established for each of the five domains of the questionnaire[22] using Item Response Theory (a psychometric modelling approach). Ceiling effects (ie area of low or no measurement sensitivity) were apparent in older typically-developing children across all domains.[22]

3.8 | Administrative burden

Most scoring involved simply summing answers to obtain a single score. Some questionnaires involved calculation of average scores for more than one domain (eg ABEL or CHAPS). The ECLiPS was a notable exception in terms of ease of scoring and interpretation. This reflects both its 5-domain structure and requirement to address age and gender effects. The test designers developed an automated scoring spreadsheet to address these complications. It means, in addition to time to enter data, access to a computer is required.

4 | DISCUSSION

4.1 | Summary of findings

Most questionnaires considered in this study had purposes at least of adequate relevance to OME and they were relatively easy to score and administer. However, publications describing the evaluation of reliability, validity, normative data, item bias and measurement model were not available for most of the questionnaires assessed. The questionnaires that scored well across the greatest number of criteria were ECLiPS, LittlEARS and PEACH. However, these three questionnaires were developed for different purposes. LittlEARS and PEACH are narrowly focused on hearing, while ECLiPS evaluates hearing and listening difficulties in the context of language, cognitive and social difficulties.

4.2 | Challenges in assessing OME

Despite clear guidelines from NICE and other bodies setting out criteria for the treatment of patients with OME,[5] clinicians frequently treat patients because of "extenuating circumstances" rather than because of meeting core criteria for ventilation tube insertion.[6] These extenuating circumstances refer to the perceived detrimental effects of recurrent infections as well as impact of OME on quality of life, speech, language psychosocial and educational development.[6] How these extenuating circumstances are established is not clear, and certainly, there is a need for an assessment tool that would allow clinicians to manage the child as a whole, not just on the basis of the pure-tone audiogram. Good quality questionnaires designed to capture these broader issues, in addition to hearing loss, could potentially meet this need.

A further challenge is one of distinguishing hearing-related problems from those linked to listening, attention and other cognitive aspects that influence a child's daily functioning. Do the children having surgery due to extenuating circumstances have problems only and predominantly due to OME? Or do these children have additional underlying language, social or cognitive problems with OME being coincidental? This is at present unknown, but reliably differentiating hearing problems from these other difficulties is clearly important, and questionnaires have a potential role to play in doing this.

Furthermore, studies on OME report a wide range of different outcome measures making comparison between them difficult.[7] The use of a well-developed psychometrically robust questionnaire which is sensitive to all issues relevant to OME would help standardise the measurement of disease impact and treatment outcome, thus aiding effective comparison between studies.

4.3 | Questionnaires in OME

Clinicians have access to many different questionnaires to support their assessment and decision-making. It can be difficult to decide which questionnaires to use. Ideally, clinicians need a questionnaire that offers information about the disease of interest that is clinically relevant, and psychometrically robust, while at the same time keeping respondent and administrator burden to a minimum. As is apparent from Table 2, no single questionnaire met all the criteria established in this study. There were considerable gaps in the formal evaluation of the quality of the available questionnaires.

As noted, when reviewing components of the analysis framework, it may have been useful to measure sensitivity of the questionnaires to treatment-related change, that is, "responsiveness." However, there is a lack of clarity on the definition of "responsiveness," meaning a lack in standardization of methodology for its calculation.[26] Furthermore, although important for interpreting changes in scores subsequent to treatment,[26] the concept provides little insight into the measurement validity of an instrument. It has consequently been argued that "responsiveness" does not need to be separately assessed when validating a questionnaire.[26]

4.4 | Limitations

Evaluation of Children's Listening and Processing Skills, LittlEARS and PEACH scored better than most of the other questions, in part because the developers included more thorough assessments of the psychometric properties of the questionnaires as part of the development process. However, these tools were not specifically designed to assess OME and may assess issues that are not strictly relevant to OME. As such, they may contain items that collect information that is redundant in the context of OME. Further exacerbating this problem, all 8 assessment criteria in the assessment framework received equal weighting in terms of relative importance. Yet, it could be argued that some components of the framework specifically conceptual clarity and validity are more important than others, such as administrative or respondent burden. The framework applied was useful for supporting a systematic assessment of the questionnaires but future researchers using this framework could consider incorporating some method for weighting the different assessment criteria according to importance relative to purpose.

In terms of item redundancy, LittlEARS was designed to evaluate age-appropriate auditory behaviour, in the pre-verbal developmental phase, while PEACH focuses on the effects of intervention with hearing aids or implants. Both of these questionnaires are narrowly focused on hearing with many items focusing on capturing symptoms of hearing loss in different contexts. This is potentially very

useful for gauging degree of hearing loss in children and infants younger than 2 years. It is less useful for older children, where the hearing related-questions are relatively redundant as clinicians can measure hearing loss reasonably reliably. They are more interested in assessing extenuating circumstances to support decisions about ventilation tube insertion. By contrast with LittlEARS and PEACH, ECLiPS offers a broad-based evaluation of listening, language, social and cognitive difficulties in children with suspected APD. Some of the items assessing these difficulties may be less relevant in the context of OME. The questionnaire also clearly misses some of the more disease-relevant issues such as impact on quality of life. On the other hand, the broader assessment of listening, language and social difficulties may help clinicians to differentiate impact from hearing loss from impact due to other underlying developmental difficulties.

It is our hope that this assessment of the various questionnaires will provide a useful framework to support clinicians in identifying an appropriate questionnaire for use in the assessment of children with OME. The methodology used in the comparison of questionnaires can be used in the assessment of other such tools used in other disease processes. The current questionnaires each have strengths and weaknesses, and our findings highlight the need for more robust assessment of measurement properties in questionnaire development. Although it may be tempting to call for development of yet another questionnaire, perhaps a better strategy would be to establish which of the existing options is most fit for purpose so that there can be a standardization of assessment procedures across practices. A variety of existing questionnaires are available, and it may be that future research or clinical assessment chooses one of the questionnaires that is most appropriate to the clinical question. On the other hand, perhaps no single questionnaire can meet all assessment needs for all OME referrals and different questionnaires may be variably informative for different subgroups or settings.

5 | CONCLUSION

We have assessed 15 commonly used questionnaires for possible use in children with OME, using set criteria. There were considerable gaps in the formal evaluation of the available questionnaires. ECLiPS, LittlEARS and PEACH appear to be the most suitable ones as they most fully satisfy the criteria laid out in this study for the assessment of questionnaires. Although initially developed for different purposes, they cover domains, which are also of relevance to OME. Questionnaires in OME could provide a useful adjunct to audiometry, as a means of assessing a child's everyday function, but further research is needed to determine how they fit into the overall assessment of children with OME.

CONFLICT OF INTERESTS

J. G. Barry is a co-author of the ECLiPS.

ORCID

R.W.C. Gan http://orcid.org/0000-0002-9383-0374
J.G. Barry http://orcid.org/0000-0002-3733-8721

REFERENCES

1. Midgley EJ, Dewey C, Pryce K, et al. The frequency of otitis media with effusion in British pre-school children: a guide for treatment. *Clin Otolaryngol Allied Sci*. 2000;25:485-491.
2. Zielhuis GA, Rach GH, Bosch A, et al. The prevalence of otitis media with effusion: a critical review of the literature. *Clin Otolaryngol Allied Sci*. 1990;15:283-288.
3. Qureishi A, Lee Y, Belfield K, et al. Update on otitis media – prevention and treatment. *Infect Drug Resist*. 2014;7:15-24.
4. Williamson I, Vennik J, Harnden A. Effect of nasal balloon autoinflation in children with otitis media with effusion in primary care: an open randomized controlled trial. *Can Med Assoc J*. 2015;187:961-969.
5. National Institute for Clinical Excellence. (2008). NICE guideline CG60: Otitis media with effusion in under 12s: surgery; available at http://guidance.nice.org.uk/CG60. Accessed March 10, 2017.
6. Daniel M, Kamani T, El-Shunnar S, et al. National Institute for Clinical Excellence guidelines on the surgical management of Otitis Media with Effusion: are they being followed and have they changed practice? *Int J Pediatr Otorhinolaryngol*. 2013;77:54-58.
7. Chessman R, Kamani T, Barry JG, et al. Which outcomes are reported by clinical trials investigating OME treatment? a case for standardised reporting. *Int J Pediatr Otorhinolaryngol*. 2016;86:93-97.
8. Williamson IG, Dunleavey J, Bain J, et al. The natural history of otitis media with effusion–a three-year study of the incidence and prevalence of abnormal tympanograms in four South West Hampshire infant and first schools. *J Laryngol Otol*. 1994;108:930-934.
9. Rosenfeld RM, Goldsmith AJ, Tetlus L. Quality of life for children with otitis media. *Arch Otolaryngol Head Neck Surg*. 1997;123:1049-1054.
10. Haggard MP, Spencer P and Multicentre Otitis Media Study Group. (2012). Short-form questionnaires can be useful clinically if powerfully derived on large scale samples. Paper presented at the Brit Soc Audiol Conference, Nottingham.
11. Andresen EM. Criteria for assessing the tools of disability outcomes research. *Arch Phys Med Rehabil*. 2000;81:S15-S20.
12. Bagatto MP, Moodie ST, Seewald RC, et al. A critical review of audiological outcome measures for infants and children. *Trends Amplif*. 2011;15:23-33.
13. Bruce I, Harman N, Wiliamson P, et al. The management of Otitis Media with Effusion in children with cleft palate (mOMEnt): a feasibility study and economic evaluation. *Health Technol Assess*. 2015;19:1-374.
14. Phelan C, Wren J. (2016). Exploring Reliability in Academic Assessment. Available at: https://www.uni.edu/chfasoa/reliabilityandvalidity.htm. Accessed March 10, 2017.
15. Antia SD, Sabers DL, Stinson MS. Validity and reliability of the classroom participation questionnaire with deaf and hard of hearing students in public schools. *J Deaf Stud Deaf Educ*. 2007;12:158-171.
16. Ching TY, Hill M. The parents' evaluation of aural/oral performance of children (PEACH) scale: normative data. *J Am Acad Audiol*. 2007;18:220-235.
17. Heidemann CH, Godballe C, Kjeldsen AD, et al. The Otitis Media-6 questionnaire: psychometric properties with emphasis on factor structure and interpretability. *Health Qual Life Outcomes*. 2013;11:201-211.
18. Barry JG, Tomlin D, Moore DR, et al. Use of questionnaire-based measures in the assessment of listening difficulties in School-aged children. *Ear Hear*. 2015;36:e300.

19. Brody R, Rosenfeld RM, Goldsmith AJ, et al. First place—resident clinical science award 1998: parents cannot detect mild hearing loss in children. *Otolaryngol Head Neck Surg*. 1999;121:681-686.
20. Purdy SC, Farrington DR, Moran CA, Chard LL, Hodgson SA. A parental questionnaire to evaluate children's Auditory Behavior in Everyday Life (ABEL). *Am J Audiol*. 2002;11:72-82.
21. Coninx F, Weichbold V, Tsiakpini L, et al. Validation of the LittlEARS® auditory questionnaire in children with normal hearing. *Int J Pediatr Otorhinolaryngol*. 2009;73:1761-1768.
22. Barry JG, Moore DR. *Evaluation of Children's Listening and Processing Skills (ECLiPS)*. London: MRC-T; 2014.
23. Kuehn-Inacker H, Weichbold V, Tsiakpini L, et al. (2003). LittlEARS Auditory Questionnaire Manual—Parent Questionnaire to Assess Auditory Behaviour in Young Children. MED-EL, Innsbruck, Austria.
24. Grimshaw SK. (1998). Assessing Hearing-impaired children's listening and processing abilities: A questionnaire and cognitive approach. PhD Thesis, University of Nottingham, MRC Institute of Hearing Research.
25. Weichbold V, Tsiakpini L, Coninx F, et al. Development of a parent questionnaire for assessment of auditory behaviour of infants up to two years of age. *Laryngorhinootologie*. 2005;84:328-334.
26. Terwee CB, Dekker FW, Wiersinga WM, et al. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. *Qual Life Res*. 2003;12:349-362.

How to cite this article: Gan RWC, Daniel M, Ridley M, Barry JG. Quality of questionnaires for the assessment of otitis media with effusion in children. *Clin Otolaryngol*. 2017;00:1–12. https://doi.org/10.1111/coa.13026