International Journal of Pediatric Otorhinolaryngology (2009) 73, 43–55

ELSEVIER

International Journal of Pediatric Otorhinolaryngology

www.elsevier.com/locate/ijporl

Air-conduction estimated from tympanometry (ACET): 2. The use of hearing level-ACET discrepancy (HAD) to determine appropriate use of bone-conduction tests in identifying permanent and mixed impairments

Mark Haggard *

MRC Multi-centre Otitis Media Study Group[1]

Received 25 May 2008; received in revised form 5 September 2008; accepted 9 September 2008
Available online 4 November 2008

KEYWORDS
Otitis media with effusion (OME);
Tympanometry;
Hearing thresholds;
Bone conduction;
Estimated hearing level;
Screening;
Mixed hearing losses

Summary

Objective: The caseload at secondary care in paediatric otology is largely otitis media with effusion (OME) and highly recurrent acute otitis media (RAOM). Few of these cases merit suspicion for hearing loss beyond the middle ear. The companion paper showed that the air conduction estimated from tympanometry (ACET) formula, derived on a very large clinical sample referred for ear or hearing problems and pre-assessed for a clinical trial, gives usable although only approximate estimates for hearing level (HL) on such a caseload. Tympanometry corresponds to a conductive loss (i.e. air–bone gap) so the HL–ACET discrepancy (HAD) should approximate the bone-conduction (BC) threshold. Clinical criteria might enable HAD to substitute for BC tests where those are infeasible, or to identify those most needing BC tests.
Method: ACET had been derived for the 4-frequency binaural average on 3085 cases with tympanometry and air-conduction HL available. On the 2701 of those with BC data at 1 kHz, we re-calculated ACET for 1 kHz only, and then explored the sensitivity/specificity trade of the discrepancy (HAD) in detecting clinically significant BC levels and the correlation between these measures. We further illustrated the generalization of the formula and cut-off on a small separate retrospective clinical sample.
Results: Correlations were moderate in the clinically relevant region. There were five cases of BC $\geq$ 30 dB in the database. At a HAD cut-off of +5 dB, the sift would identify

Abbreviations: OME, otitis media with effusion; HL, hearing level; AC, air conduction; BC, bone conduction; PCHI, permanent childhood hearing impairment; ACET, air conduction thresholds estimated from tympanometry; HAD, hearing level-ACET discrepancy.

* Corresponding author at: Psychological Laboratory, Downing St, Cambridge CB2 3EB, United Kingdom. Tel.: +44 1223 363609; fax: +44 1223 335654.
E-mail address: mph38@cam.ac.uk.

[1] Full list of authors and contributors are given in Appendix B; Guarantor: Professor Mark Haggard; Writing and analysis team: Haggard MP, Higson JM, Spencer H.

0165-5876/$ – see front matter
doi:10.1016/j.ijporl.2008.09.023

all (nominal 100% sensitivity). For marginal cases, two definitions were adopted (BC $\geq$ 25 dB and $\geq$20 dB; 9 and 23 cases, respectively). Sift sensitivity remained high (89% and 83%, respectively), and specificity was acceptable (75% for both definitions). *Conclusions:* Given tympanometry and air-conduction HL, comparison of HAD with a recommended cut-off gives acceptable sensitivity and specificity for non-OME hearing problems. BC testing can be reserved for probable positive cases, provisionally only 25% of caseload. HAD could temporarily substitute for BC measurement in children too young to accept bone-conduction transducers in awake testing. Where a high proportion of the caseload is expected to have middle ear fluid, ACET and HAD together offer efficient possibilities for best use of available information.

1. Introduction

Bone-conduction (BC) audiometry is a key test in the confirmation of suspected permanent childhood hearing impairment (PCHI—i.e. sensorineural hearing impairment, or permanent conductive abnormalities not due to middle ear fluid). Conventional air-conduction (AC) audiometry measures the sensitivity of the entire auditory system, but BC largely bypasses the structures of the middle ear, thus assessing the integrity of the cochlea and cochlear nerve. The role of these tests may be affected by the recently introduced universal neonatal hearing screening systems (UNHS). These yield about 1 PCHI case per 1000 infants tested [1], concordant with reported incidence 0.91—1.07 per 1000 [2], depending on case definition. However not all cases can be detected at birth; due to imperfect sensitivity. Acquired and progressive cases will arrive at paediatric ORL/audiology clinics, where over half the work of diagnostic suspicion and confirmation will remain, by a non-screen route. Fortnum et al. [2] reported an almost doubling of the cumulative prevalence of identified PCHI from 3-year-olds to 9—16-year-olds, as these cases become identified. Hence in a cohort experiencing UNHS, the number of early cases of PCHI will be reduced but the caseload of later referrals will not change much. Finding these remaining PCHI cases efficiently poses a challenge to ORL as greater vigilance is required for rarer conditions. The majority of the paediatric otology and audiology caseloads present with otitis media with effusion (OME) [3]. The proportion presenting also with PCHI (largely sensorineural), either alone or in combination with OME can be expected to be higher than seen in the general population as suspicion and referral are of some validity even if this proportion may be insufficient to support an early intervention policy. The combination of co-present conductive and sensorineural hearing loss, even if each is mild, may precipitate referral of a case, which either loss existing alone may not.

The typical age for non-screen referrals raises feasibility problems. The earphones for conventional AC measures are usually well tolerated in children and, with appropriate play and reinforcement techniques, can give accurate results for young infants. However if a BC test is to be well done, with its tight headband and often coming at the end of a testing session, it can consume much clinic time and much of the younger child's tolerance [4]. Consequently, some clinics choose to carry out BC only where they consider it "necessary", but the basis of this necessity is not well defined. We recently surveyed a sample of UK audiology clinics on their practice for testing children with BC [5]. Of 50 respondents, 64% declared a policy of testing *all* paediatric otology children with BC. The remaining 36% varied widely in the proportions which they claimed received a BC test, and in what their criteria were for giving it. The declared bases for receiving a test included the history, shape of AC audiogram and child's age and cooperation; wide differences were reported in use of masking and in number of frequencies tested (from 1 to 7). Declared policies are often not followed because of clinical pressures, so the realities may be less impressive than the claims, but the point is that such differences even in the claims reveal considerable variation in clinical practice. Some practice variation over time and place is to be expected, particularly in a system that is driven by an uneasy balance between three factors other than precisely determined information needs in individual cases: consensual (if not always evidence-based) clinical best practice, resource availability and the tolerance or concentration of the child [6,7]. However there is also a professional and scientific responsibility to minimize practice variation, as it rarely has a justified basis, and a high degree of variation is a marker of low overall system effectiveness.

In particular clinics, the local practice may translate into inefficiencies (more testing done than is productive), which are costly in monetary terms; or it may translate into insensitivity (too few BC tests done), which has a penalty in terms of cases missed. Where decisions are not made systematically, the

process may oscillate randomly between these two extremes. Implementing a protocol to determine when and for whom BC tests should be performed would reduce practice variation, and hence improve the success of detecting PCHI cases. For those selected, better/fuller BC testing would involve attending to careful placement of the bone vibrator (possibly including a re-positioned replicate) and including an appropriate number of frequencies. This could be a better use of resources overall than doing one or two determinations on all cases (leading to repeat tests or possibly missing/underspecifying true PCHI).

Uncertainty about the need for BC does not arise in more severe cases (AC $>$ 50 dB) for which a purely OME explanation becomes near-impossible. In such cases BC testing is clearly indicated. Similarly, children with normal AC thresholds will not require BC tests. However, the many children with AC thresholds between about 20 and 50 dB might have a mixed loss with 5% to 95% of it being PCHI. Although the balance of probabilities will always favor a diagnosis of uncomplicated OME, we propose a sift (case triage) to further reduce this uncertainty in an efficient way. Although the essentials are the same, the term ''screen'' is not appropriate for a contingent, within-clinic process.

2. Evidence base for deciding who should receive bone-conduction testing

To address the issue of efficient identification of mild to moderate PCHI within a predominantly OME caseload, we have used cases assessed for TARGET, a large multi-centre randomized trial in OME [8—11]. Baseline TARGET data provided 23 cases with BC thresholds at or above 20 dB, within a sample of 2701 cases who had both a BC test result and an air conduction estimated from tympanometry (ACET) score. The companion article [12] describes the derivation of the ACET formula. This formula was obtained by multiple regressions, configured to best predict the true binaural average four-frequency hearing level (HL) from binaural tympanogram categories. The ACET score has other applications, but we here develop the one that relates to BC testing. By providing a justified and moderately predictive estimate of the hearing from the child's tympanometric profile, ACET broadly quantifies the conductive component due to OME. This is essential background to considering whether some part of the HL (AC) may be due to PCHI. Thus, when the HL—ACET discrepancy (HAD) exceeds a certain amount, the HL is ''unexplained'', and BC testing would be indicated. Many good clinicians follow similar reasoning informally, comparing HL to the apparent severity of OME, and calling for BC testing if HL seems disproportionate. The HAD criterion corresponds to such a clinical rationale, but does so (a) consistently over occasions, (b) optimally and precisely, so performing well in non-expert hands, and (c) with documentable accuracy and reliability. Its workings are intuitively comprehensible, because tympanometry coded as ACET can be considered conceptually equivalent to an ''air—bone gap'' (i.e. the difference between HL measured by AC and BC). It is convenient for comprehension and presentation that the four measures in the steps of the argument are on scales with the same meaning and units—dB HL. They are: (i) the prediction from tympanometry (ACET); (ii) the true measured HL, (iii) the difference between these (HAD); and (iv) the true BC.

Some cut-off for general importance or specific actions has to be built into all clinical decision procedures, generating small errors in the process. Errors have to be minimized in a way that maximizes clinical value. True-positive and false-negative decisions determine the sensitivity and specificity of a screen or sift, so these are documented for a clinical algorithm routing cases to BC testing. Adopting such an explicit algorithm would serve to minimize practice variation.

3. Method

3.1. Samples

Two samples are used in this paper:

(1) TARGET sample (children from the baseline recruitment phase of the TARGET trial having the necessary data).

The first assessment visit in the trial of alternative regimes in glue ear treatment (TARGET) study provided the derivation sample. TARGET is a trial of treatment outcomes in OME and followed 432 children (376 randomized, 56 more severe cases treated) for 2+ years through treatments of ventilation tube insertion with or without adjuvant adenoidectomy versus watchful waiting. The database also addresses many epidemiology and follow-up issues beyond the 2-year trial outcomes. For the present analyses, the sample available consists of 2701 children meeting all the following requirements: aged $3^1/_4$ to $6^3/_4$ years, with complete air conduction (AC) thresholds (0.5—4 kHz), an unmasked bone-conduction (BC) threshold at 1 kHz and

an interpretable tympanogram recording for each ear. These children are a subset of 4816 documented referrals, with a history of ear or hearing problems but no previous ORL operation, monitored for the TARGET trial at one of 14 ORL centres in the UK over a 4-year period in the late 1990s. For children not qualifying for TARGET, centres were not required to supply audiometry results, but it was the routine practice of most centres to acquire such data, yielding 2701 cases with complete data for the variables listed. Approximately equal numbers of children were referred from family practitioners and community audiology services. The generalizability and properties of the TARGET sample are described elsewhere [10] and in the companion paper [12].

(2) Warrington clinic sample.

To test and illustrate the applicability of the procedure to a real caseload, an anonymized opportunistic clinic sample (22 children similarly aged but audiologically varied for illustrative purposes) was composed from routine referrals to community audiology services in Warrington, UK, for ear or hearing problems. In the UK such paediatric community audiology clinics have a diverse role and receive children suspected for hearing loss, language and behaviour problems, but not necessarily suspect for OME. Children were aged 3 years to 7 years, 5 months and had complete AC thresholds (0.5—4 kHz) and an interpretable tympanogram recording for each ear. Complete unmasked BC thresholds at 0.5—4 kHz were available for 19 children and the remaining 3 had at least 3 of these thresholds, always including 1 and 2 kHz. These supplementary cases are referred to as ''clinic'' children, not being in a trial.

4. Measures

4.1. Audiometry

In both samples, audiometric thresholds at 0.5, 1, 2 and 4 kHz were obtained for each ear according to method A of BSA recommended procedures [13]. Play audiometry techniques were an option for younger children, generally those less than 5 years [9]. For no child was AC asymmetry or the air—bone gap great enough to indicate masking of AC thresholds. BC thresholds were also unmasked, for the practical reason that BC testing is generally poorly tolerated in many young children, making the full determination of masking levels infeasible for a large study. As there is little or no transcranial transmission loss [14], and effects of transducer placement are local and not systematic, BC data apply to the better-hearing ear.

4.2. Tympanometry

Tympanometry was performed with sweeps at 50 daPa/s down to −400 daPa (and for some flat traces in the TARGET sample with a further sweep down to −600 daPa). Outputs were categorized according to the modified Jerger scheme [15,16]. Raw data were entered on an Access database, and displayed and analysed with SPSS Version 12.

4.3. Calculated measures

4.3.1. ACET

ACET was based on estimating *binaural* average HL (0.5—4 kHz) from tympanometry on each ear. It maximises the accuracy of estimates through aggregating reliability over frequencies and ears, and by exploiting a demonstrated inter-dependence (contralateral conditioning) between the ears (see companion paper [12]). Here we derive and use a slightly different version of ACET to match the limited BC data available (1 kHz). The HAD principle connects the thresholds at corresponding frequencies in AC and BC, and so it was important to avoid confounding this principle with unnecessary error variance due to differences in frequency. Thus the present version of ACET predicts the better-ear HL at 1 kHz from binaural tympanometry (see Appendix A); the model includes the contralateral conditioning effect of the other ear's tympanogram status (i.e. there was a significant binaural interaction) and explains 44.8% of the total variance at 1 kHz. The general justification for preferring a comprehensive reduction of *binaural* tympanometry and for incorporating an interaction (multiplicative term) in the preferred formula are documented in the companion paper.

4.3.2. HL—ACET discrepancy

For the TARGET dataset, HAD values are adjusted differences: the saved residuals (deviations in HL) from the regression line between observed (HL = y) and predicted (ACET = x) of individuals' data points. For the clinical dataset, HAD is a simple subtraction of the ACET score, calculated from the regression formula (see Appendix B), from the true 1-kHz better-ear threshold. A discrepancy (HAD) of +5 means that the obtained average AC threshold is 5 dB worse than that predicted by our formula from the binaural tympanometric configuration. This need not correspond exactly to 5 dB of sensorineural loss (BC) because of the additive constants

in regressions, although the data show that the two measures (HAD and BC) are in fact quite close.

5. Results

5.1. Characteristics of samples

Table 1 gives the profile of 23 TARGET and 5 clinic children having a BC value ≥20 dB, and is ordered by BC threshold and then age. The mean thresholds are significantly higher in the clinic sample at 4 kHz, although not at 1 kHz, consistent with there being more PCHI cases in that group (see column 2). The mean HAD values were significantly higher in the clinic group, as expected: more of the HL in that sample is PCHI, so is not explained by tympanogram type.

Table 2 cross-tabulates for the TARGET sample, BC severity with presence/absence of a B-tympanogram on at least one ear. In this section only, the term "sensorineural" is used in preference to PCHI, because cases are defined on the basis of BC. Use of the PCHI definition elsewhere retains the possibility of conductive losses other than those (mostly OME) due to tympanic membrane immobility. The definite cases (30 dB+) in this contrast did not differ from the marginal ones in age, nor in the proportion arriving

Table 1 Audio-tympanometric profile of cases with PCHI, as detected by BC thresholds at or above 20 dB, from (a) TARGET (N = 23) and (b) clinic (N = 5) cases.

Age (months)	Bone-conduction (BC) threshold	Air conduction (AC) thresholds			Tymp (left)	Tymp (right)	HL—ACET discrepancy (HAD)
	1 kHz	1 kHz (better ear)	4 kHz (left)	4 kHz (right)			
(a) From the 2701 TARGET cases with complete audio-tympanometry and age in range							
44	20	40	45	45	B	B	8.97
45	20	50	60	60	B	B	19.05
46	20	40	50	50	B	B	9.14
47	**20**	**20**	**50**	**30**	**C2**	**C2**	**−0.29**
52	20	40	15	35	A	B	24.19
53	**20**	**20**	**30**	**25**	**B**	**B**	**−10.24**
55	20	45	35	30	B	B	14.93
56	20	40	25	60	B	B	10.02
60	**20**	**25**	**50**	**40**	**B**	**B**	**−4.63**
69	20	50	50	50	B	B	21.15
71	20	50	40	35	A	A	37.95
78	20	40	35	50	B	B	11.94
78	20	40	40	40	A	A	28.56
81	20	20	15	30	C1	A	8.83
48	25	45	50	45	C2	B	24.42
54	25	25	15	10	C2	A	9.58
62	**25**	**10**	**45**	**15**	**A**	**B**	**−4.93**
78	25	50	20	25	B	B	21.94
52	30	40	25	20	A	A	26.29
64	30	40	45	40	A	A	27.34
69	30	30	15	5	A	C1	17.78
57	45	50	50	45	A	A	36.73
81	60	50	45	30	A	A	38.83
Means							
60.86	25	37.39	36.96	35.43			16.42
(b) From the 22 clinic cases with complete audio-tympanometry and age in range							
72	20	30	45	50	C2	C2	13.00
84	35	45	60	50	A	A	34.99
82	40	60	85	85	B	B	32.37
89	45	45	85	80	A	A	35.39
84	50	40	75	80	A	A	29.99
Means							
82.2	38	44	70	69			29.15

Cases in bold type would not be detected by the +5 dB cut-off in HAD (the discrepancy between true and tympanometrically predicted HL).

Table 2 Cross-tabulation of mixed/pure basis of hearing loss with BC threshold in cases identified with sample of referrals in TARGET baseline data.

BC stratum	Type B tympanogram on at least one ear		
	Yes	No	Total
20–25 dB	13	5	18
30 dB or above	0	5	5

Fisher's exact probability test for 2 × 2 association: $p = 0.007$.

at ORL via the two main routes. These routes are a direct one from family practitioners (FPs) and an indirect one involving referral through community audiology (CA) services. Although not in fact relevant, this last distinction could have been, because prior audiometry — available at CA but generally not at FP — improves both sensitivity and specificity of referrals.

The cross-tabulation suggests strongly that the majority (13/18) of the cases with marginal BC have entered the sample of referrals because their overall hearing disability has been increased by an OME conductive loss. The five definite cases (BC $\geq$ 30 dB) show up without a conductive loss, and this contrast in pattern is significant (Fisher exact $p = 0.007$). The binaural four-frequency average hearing levels (AC) for these two groups are very similar: (37.7 dB, S.D. 11.0 dB for 18 mild cases, mainly mixed with marginal sensorineural underlay, versus 39.75 dB, S.D. 10.5 dB for the five definite sensorineurals). This similarity suggests chiefly a hearing disability basis for referrals (whether informal or formal), and that this is closely related to total dB HL. On the other hand, the usefulness and validity of the HAD measure is demonstrated by the mean difference in HAD, between the marginal mixed and distinct sensorineural cases. The group difference in mean HAD is 16.58 dB (+12.81 dB, S.D. 12.66 dB for mild/mixed cases, versus +29.39 dB, S.D. 8.54 dB for distinct sensorineural cases; Mann–Whitney $U = 12$; N1 = 5, N2 = 18; $p = 0.012$). In agreement, the five clinic children with raised BC had a mean HAD score of +29.15 dB (S.D. 9.29 dB), very close to that of the five definite (i.e. pure, not mixed) sensorineural cases within the TARGET sample. Indeed, four out of five of these clinic children have BC > 30 dB, but tympanograms indicate that one of the four also has a mixed loss. The HAD-values show the relevant clinic children to be similar to the definite sensorineural cases within the TARGET sample.

5.2. Relationship between HAD and BC

Fig. 1 shows, in the dataset combining the TARGET and clinic cases, the relationship between the BC threshold and HAD. Taken over the total TARGET sample, the correlation between these two variables, although significant on the very large sample, was not high ($r = 0.171$). The entire dataset does not offer the most relevant way to address the ability of HAD to predict BC, because the sample is swamped by a large number of cases with normal BC and normal HAD, among whom there are other sources of variance in these measures, making them only weakly related. The clinic sample appears to contrast in this respect, having a much higher correlation of BC with HAD for its 22 cases ($r = 0.581$). Usually, restricting the range in one or both variables *reduces* a correlation's magnitude. However, here restricting the TARGET sample to the tail of cases with BC $\geq$ 15 dB makes it more like the clinic sample from a descriptive point of view. This similarly *improves* the correlation from 0.171 to 0.494 ($N = 60$, $p < 0.001$), with a further improvement to $r = 0.536$ ($p < 0.001$) if we further restrict to the 23 cases with BC $\geq$ 20 dB. Thus the correlation in the practice-relevant range — the tail of cases with at least some marginal problem — is much stronger than it is overall. These correlations confirm that in general, higher HAD scores are moderately predictive of BC abnormality, so they can indicate the need for BC testing.

5.3. Setting an appropriate criterion

The stronger relationship between HAD and BC over the higher range of BC values could underpin a protocol whereby BC testing is performed only on those with sufficiently high HAD scores. The major issue for application is: how high a cut-off in HAD is necessary to find a high enough proportion of the PCHI cases? From Fig. 1 we can see that choosing a

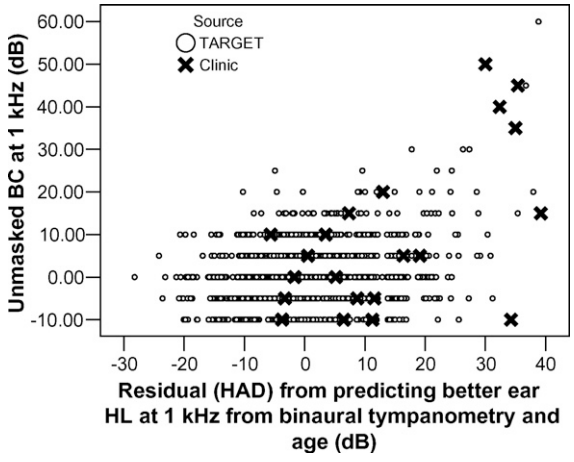

Fig. 1 BC threshold plotted against residuals from regression predicting better ear HL at 1 kHz from three-category tympanometry and age ($n = 2701$).

HAD cut-off anywhere between 0 and 8 dB, would ensure that all but four cases with BC $\geq$ 20 dB (Table 1) would be found. Of the four cases *not* so found, three have BC at exactly 20 dB and the fourth, although with BC = 25 dB, is also likely to be marginal by virtue of the air conduction threshold being 10 dB. These four mis-classifications could therefore be considered acceptable. However, without more true cases in the dataset between HAD values of about 0 and 8 dB, a cut-off at any value in this range will give a similar estimate for sensitivity and a more precise optimum cannot yet be specified.

The inevitable sensitivity/specificity trading relationship in any screening or sift application forces an explicit valuation of the benefits and disbenefits of the various types of error resulting from setting a particular cut-off. The standard representation for examining these issues is the Receiver operating characteristic (ROC) curve. Fig. 2 shows the ROC for detecting cases with BC $\geq$ 20 dB for differing levels of HAD in the TARGET data. The ROC points are calculated for 2 dB steps in HAD, ranging from −10 to +40 dB. The plateau of 82.6% sensitivity is seen here for a range of specificities (depending where in the 0—8 dB region the cut-off is placed).

The most appropriate cut-off value combines maximum sensitivity with acceptable specificity, hence acceptable follow-up effort (here the percentage requiring BC tests). Table 3 shows the usual sensitivity/specificity trade-off for differing choices of HAD cut-off to detect BC $\geq$ 20 dB and also $\geq$15 dB and $\geq$25 dB. If detection only of 25 dB losses and above is required, the sensitivity improves to 88.9%. Using a cut-off of +5 dB in HAD, only 25% of the caseload would require further testing. If a more severe definition is set for cases to be found (30 dB BC or worse), then the +5 dB HAD cut-off achieves 100% sensitivity (with still only 25% of cases requiring testing). However, to detect milder HLs (BC) at 15 dB or above, a lower HAD cut-off, leading to a requirement for more BC testing would be required. For this case definition, proceeding to test BC on 34.0% of the caseload would be required to achieve 73.3% sensitivity. The data thus already support a range of acceptable and efficient scenarios in the likely zone for optimal cut-offs. We do not claim that the central scenario to date ($\geq$20 dB, HAD cut-off +5 dB or higher) will be absolutely ''right'' for all circumstances.

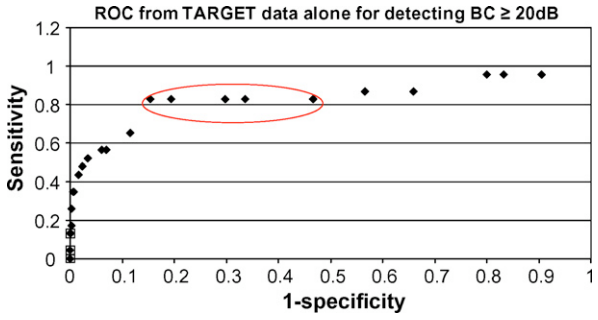

Fig. 2 ROC for detecting a BC threshold $\geq$ 20 dB for various cut-offs in discrepancy between the tympanometrically estimated binaural HL and true binaural HL at 1 kHz. The circle defines the range of unresolved sensitivity due to low case frequency (see text).

Table 3 Sensitivity, specificity and percentage of caseload requiring to be tested with BC, to detect BC hearing levels at or above 15, 20 and 25 dB, for given HAD cut-offs (TARGET baseline data).

To detect	HAD cut-off (residual)	Sensitivity	Specificity	% to be tested
$\geq$15 dB	**2** [a]	73.33%	66.91%	33.99%
	6 [b]	60.00%	81.03%	19.88%
$\geq$20 dB	**5** [a]	82.61%	75.32%	25.18%
	8 [b]	82.61%	84.62%	15.96%
$\geq$25 dB	**5** [a]	88.89%	75.04%	25.18%
	8 [b]	88.89%	84.29%	15.96%

[a] Values in bold accept a slightly poorer specificity than those giving optimum prediction according to the ROC curves, in order to be sure to capture relevant cases needing bone-conduction testing in datasets containing similar referrals but inevitably differing slightly.
[b] Value giving rise to maximum sum of sensitivity and specificity.

5.4. Applying cut-offs from derivation sample to clinic data

Recalculating the ROC curve and Table 3 for the combined data sets did not change the sensitivity/specificity trade nor add precision to the sensitivity estimate (data not shown). The clinic sample also had very few cases in the marginal zone where they could make a difference. We therefore here use the clinic data chiefly to illustrate the application of cut-offs proposed from the TARGET data. Choosing a +5 dB cut-off in HAD would in the clinic sample lead to 68% (of those children with a priori a higher percentage of mixed and PCHI problems) requiring BC tests, and would detect 100% of cases at above 15 dB of BC. A higher cut-off of +8 dB in HAD would in this sample have reduced the testing percentage to just over half, and would still have achieved 100% sensitivity for cases of 20 dB BC and above (85.7% sensitivity for 15 dB and above). If the clinic sample were itself to be taken as the reference source in setting criteria, a much higher cut-off (up to 29 dB in HAD) would be implied, detecting 80% of cases at 20 dB BC and above and 100% of cases 25 dB and

above. However this is not an appropriate basis for fixing a clinical cut-off, because the supplementary clinic sample is not a general ORL otology caseload within which the present problem of detecting fairly rare PCHI and mixed cases is defined, and where it is perhaps most needed. Furthermore the clinic cases are somewhat selected for diversity in illustration, so the sample is not even representative for paediatric community audiology clinics. This observation leads to two cautions: (a) detection of the PCHI cases in the clinic sample can be achieved with a higher HAD cut-off, so a cut-off optimised on one type of caseload may not automatically transfer to another; (b) the milder PCHI and mixed cases are seen chiefly in the TARGET sample, based on referrals to ORL for OME, and referred for hearing loss temporarily made worse by OME (see Table 2).

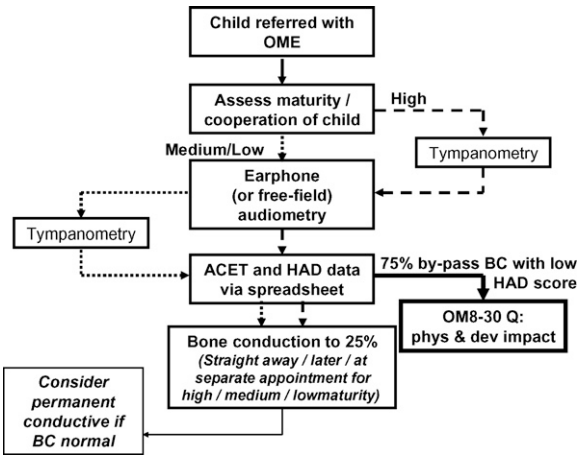

Fig. 3 Suggested clinical testing algorithm for handling possible overlap of OME and PCHI.

6. General discussion

6.1. Choice of cut-off

The choice of cut-off for HAD that would indicate need for BC testing has implications for clinic effort. It is important not to cherry-pick over-favorable cut-offs (such as the example of +8 dB in Table 3 which suggested that only about 14% would need to be tested). We have therefore provisionally accepted a cut-off value (i.e. +5 dB) that is middling for the set of cases giving the sensitivity plateau; the slightly lesser specificity thus achieved corresponds to a larger but more realistic testing burden. The +5 dB value for optimum cut-off gives a BC testing burden of 25% for detecting 83% of cases having 20 dB BC thresholds, or detecting 89% having 25 dB or higher thresholds. Such an integrating and non-opportunistic choice of cut-off is more likely to be replicable on other datasets having more cases in this region, than a +8 dB cut-off (so more attractive in resource terms) would be. A HAD cut-off of +8 dB would make the option of avoiding BC tests seem more efficient so is attractive, but the price is increased possibility of missing a case with 20—25 dB BC. We therefore recommend within the range of maximum sensitivities accepting a central value of +5 dB for good balance of sensitivity with specificity, plus a realistically achievable reduction in testing burden.

6.2. Clinical pathway

Fig. 3 suggests a pathway that a clinician could follow to use the above findings in assessment of OME referrals. Use of a spreadsheet (see later) means that ACET and HAD calculations can be done in real time and not greatly delay the patient pathway through the clinic or add to overall clinic time. Table 4 gives a worked example of how the pathway would be followed. The figure also considers the practical issue of child concentration and co-operation. Tympanometry is usually well-tolerated in children, with many commercial tympanometers operating child-friendly visual outputs. However some younger children with limited duration of attention or co-operation may benefit from a sequence which obtains the crucial AC thresholds first. Lost co-operation or lost concentration after tympanometry may then not matter. For a very young or immature child, free-field audiometry would be appropriate. For such a child, were he/she one of the 25% for whom BC is indicated, this test would have to be deferred, and usually can be, to a time when the child is older. However, in the meantime, the algorithm can via tympanometry and air-conduction HLs produce a surrogate BC measure, in the form of HAD, as an evidence-based working clinical hypothesis. Further work is needed on the specification of free-field testing procedures and calibration to fully apply the ACET principle to them.

In the companion paper [12], a trichotomous version of ACET was also developed predicting the four-frequency binaural average HL from a three-level binaural tympanogram variable (0, 1 or 2 type B traces) and age. This simplified model for clinical use without computer support gave very similar predicted HL values (to within 1 dB) to those which the original ACET score predicted. Its output is one of three stereotype predicted ACET values adjustable by ± 1 dB for age. Given the further precise comparison required with true HL, the need for decimal arithmetic and the need for a traceable account of decisions taken and their reasons, we

Table 4 Worked example of assessment for a real case (grey boxes would be computed by the spreadsheet).

Presentation

Child A, 5 years 4 months, is referred with hearing problems:

Preliminary hearing assessment gives:

- Bilateral A tympanograms.
- **HL at 1 kHz in better ear: 40 dB**.

Algorithm

Step 1:	Enter bilateral tympanogram types (A, A) and age (64 months) into Excel™ spreadsheet to produce ACET.	ACET = 12.7 dB
Step 2:	Calculate HL-ACET discrepancy (HAD) (HL-ACET) discrepancy = 40 - 12.7 dB	HAD = 27.3 dB
Step 3:	Compare HAD with cut-off (+5 dB). 27.3 dB is much higher than cut-off.	BC testing strongly indicated
Step 4:	Obtain average unmasked BC threshold*	Unmasked BC = 30dB Child's SNHL confirmed.

Summary

The HL-ACET discrepancy (HAD) indicated that tympanometry did not explain the child's hearing loss and so BC testing was needed. In this example, the HAD predicted the unmasked BC threshold particularly well.

have assumed use of the spreadsheet rather than a look-up table of age-adjusted ACETs, so have not applied the trichotomous model to this clinical problem. A +5 dB cut-off would be equally appropriate for the trichotomous model if this were used without computer support.

Of the cases delimited by the HAD cut-off (i.e. those with HL not fully explained by tympanometry via ACET), as well as the true sensorineural cases and some false positives, some types of permanent conductive case could potentially be identified. This would be because, as for the sensorineural cases, their AC sensitivity is not sufficiently explained by the tympanogram. As BC test results will also be mostly normal for these cases, it is important that they not be confused with the false-positives. Fully solving that clinical problem is beyond the present scope, but in practice this is not a great drawback: many of such cases will have been identified previously from craniofacial examination, and the data will still indicate an unexplained AC hearing loss requiring consideration.

6.3. Unilateral and bilateral issues

Asymmetrical PCHI poses a separate type of issue. Clearly a binaural procedure able to identify 20 dB binaural BC loss could only identify approximately 40 dB unilateral loss. A case with such a mild unilateral loss would generally not receive a hearing aid. Despite a biomedical tradition of wishing to measure status of each ear as precisely as is feasible, we regard the lessened test sensitivity for unilaterals as acceptable. Within this pragmatic binaural whole-child approach, time saved by following an algorithm such as ours can be usefully redeployed, for example in an OME diagnosis to make a systematic assessment of severity of impact and wider clinical implications with questionnaires [17]. To facilitate use of ACET and HAD scores in routine practice, clinical audit and research settings, we have produced a spreadsheet enabling data entry, storage and transmission. The spreadsheet additionally makes imputations to allow for formula-based values to be generated, even if the full data (e.g. all four audiometric frequencies on each ear) are not available. It does so in a way that takes account of the average OME audiogram. A user group of those interested in the potential gains for appropriate and economical clinical decision making is being established to acquire further data to test and extend ACET applications such as HAD.

6.4. Validity of the 1-kHz test frequency used

The test frequency used here (1 kHz) may represent the issue of BC losses in a way that safeguards against over-medicalizing the mild condition of hearing loss restricted to high frequencies. The

prevalence and incidence would be higher in a population study if cases with hearing loss only in high frequencies were identified. However, in our trial sample (both the complete sample and those with abnormal BC) the 4-kHz threshold in each ear is from 3 to 5 dB better than at 1 kHz, because 1 kHz is more affected by the OME. Even though mild late-detected sensorineural losses could be expected to have chiefly the high frequencies affected, this seems not to be true of the present cases. This fact makes the case-finding for mild sensorineural hearing loss not primarily a matter of high-frequency testing. The single frequency of BC testing available here (1 kHz) is a shortcoming of the data used here but not a compromising one.

6.5. Definition of a case worth finding

We do not claim to know whether, irrespective of conductive overlay, 20 dB of PCHI is worth finding early; no randomized trial withholding intervention on a sample free of the selection biases operating in early identification in the past has been reported addressing this issue. However there is some *prima facie* case on grounds of impact on schooling [18,19] that mild and unilateral cases do meet problems, and hence might be worth finding early. In the long term, having a correct pathophysiological explanation (e.g. a mild hearing loss) of any difficulties met should be preferred to having an incorrect psychological one (e.g. inattentiveness, low intelligence), provided that stigmatization, over-reaction, over-dependency and over-intervention are avoided.

Research on a wider basis would be required to conclude on what scale OME brings milder PCHI to light in practice, via addition to form a non-trivial degree of mixed loss; our results merely show that it can. However the additive principle and the present findings on mixed losses should orient audiologists and otolaryngologists post-UNHS to this as an example of compound pathology that they need to be on the look out for as contribution to case-finding. Marginal losses of each type in isolation will generally not turn up in the caseload. But the fact that mixed losses, in which each component may be marginal, do turn up underlies the additivity of hearing loss and the probable need for management of mixed losses. The difference of only a few dB between the AC thresholds of the PCHI and the mixed loss cases within the present sample is consistent with the rough equivalence [20] of impact from the two forms of hearing loss, and hence with the usefulness of dB HL impairment as a surrogate marker for disability.

6.6. Appropriate management for mild and mixed hearing losses

The appropriate management of the mixed cases that would be identified by the sifting and confirming processes suggested here has two aspects—comorbidity and range for conferring material benefit. It is widely believed that an underlying comorbid condition should influence the approach to management, making it appropriate to err in the direction of intervention. Here each type of hearing loss can be considered a comorbidity for the other. A cogent general case has been put [21] for considering the many comorbidities and vulnerabilities to OM(E) impact in this way, even though management of all OME cases is becoming more conservative than in the past. Controlled trials are unlikely to be undertaken on such a small clinical group as the mixed cases spotlighted here, so rational principles and case series are needed to arrive at an overall clinical strategy for them.

We have chiefly defined cases as $\geq$20 dB BC in the present analysis for the following reason. In pure OME in mid-childhood, the average measured AC hearing loss after intervention with ventilation tubes (VTs) is not 0 dB, but about 10—12 dB, similar to that in children referred but currently not showing middle-ear fluid as marked by B tympanograms [22]. Thus in a child with a mixed loss having a permanent 20 dB BC component, VT placement for OME might reduce the AC loss from over 50 dB down to 30—32 dB HL, still material. The obvious clinical issue arises of counselling to expect some definite improvement with VT placement in such a case, but also to limit such expectations. It is also possible to consider the additional use of a hearing aid, even with ventilation tubes, until the air—bone gap eventually returns nearer to zero. For a BC threshold of only 15—20 dB, managing the remaining mixed loss including the sensorineural component with aid(s) is more problematic; repeat measures, a scrutiny of the frequency pattern, and a distinction between the short- and long-term strategies would be required.

7. Conclusions

Where otitis media with effusion combines with permanent hearing impairment, the summed impairments can exceed a criterion for communication problems and concern, compared to children with either condition alone. Thus OME raises the probability of referral and identification of cases where the permanent component alone could have gone un-noticed. Issues of diagnostic separation and short- and longer-term management arise.

Despite the need to note comorbidities in OME, many cases do not need BC testing. The basic audiometric assessment of presumptively OM(E) children (air-conduction audiometry and tympanometry) can be combined into a predictive sifting criterion to locate probable cases of material PCHI or mixed loss. We offer a method to do this, based on ACET — the air conduction threshold estimated from tympanometry — as introduced in the companion paper.

By applying an appropriate criterion to the discrepancy between the tympanometry-based ACET and the true HL measure but AC, confirmatory BC tests need only be given to a proportion of OM(E) cases. This is provisionally 25% at a discrepancy of true HL being +5 dB worse than predicted. Sensitivity of this sift is acceptable, with only marginal BC values being missed.

The HL—ACET discrepancy gives a basis for selectively introducing BC testing in clinics where BC is *not* at present done, for re-distributing effort more efficiently where BC is *usually* done and for adopting a more rigorous, precise and evidence-based criterion for testing where it is *sometimes* done.

Population screens of hearing later than neonatally may become difficult to justify on the basis of small yield, so primary and secondary healthcare need to collaborate to optimise an overall system for vigilant surveillance, with cost-effective referral and assessment. Within this, the availability of a predictive technique to contain the amount of BC testing in referred cases will help to keep the implied clinic workloads manageable.

Acknowledgements

We acknowledge the contribution of all the hospital chief executives, medical directors, clinical managers and pharmacists in facilitating the trial.

Appendix A. The formula for air conduction HL estimated from tympanometry (ACET) when better ear threshold at 1 kHz is predicted

The methodology for predicting HL from tympanometry is given in the companion paper [12]. We here give a modified ACET formula where the variable predicted in the regression is the 1 kHz better ear (BE) threshold. This version of ACET was appropriate for comparison with BC which was available at 1 kHz only. As in the other ACET derivations, the model here includes interaction terms expressing the average contribution of the tympanogram status from each ear on predicted thresholds. The model coefficients and examples of application are given below. The present formula explains 44.8% of the variance in better ear HL at 1 kHz, compared with 49% explained in predicting average four-frequency HL reported previously, the drop being most simply explicable as lower reliability for one frequency measure than for the average of eight.

B-coefficients specifying the regression formula for predicting better ear hearing threshold at 1 kHz.

Regression constant	Categorical variables					Linear variable
	Main effects			Interactions		Age (months)
	Tympanogram types	Left ear	Right ear	Left × right tympanogram combinations		
34.9	A/C1	−14.5	−14.1	Left A/C1 × right A/C1	+12.0	−0.1
	C2	−10.1	−8.8	Left A/C1 × right C2	+8.2	
	B	0	0	Left C2 × right A/C1	+9.4	
				Left C2 × right C2	+8.4	
				At least one ear type B	0	

To apply the categorical variables, one number for each ear is selected from the "main effects" field, corresponding to the tympanogram types on each ear; one number is selected from the "interaction" field corresponding to the pairing across the ears. To apply the linear variable (age), the child's age (in months) is multiplied by the coefficient given. The predicted better ear HL is the regression constant plus these four signed contributions.

Example: A child of 5 years (60 months) has a left-ear A tympanogram and a right-ear C1 tympanogram:

$$\text{Predicted better ear HL at 1 kHz} = 34.9 + (-14.5) + (-14.1) + 12.0 + (-0.1 \times 60) = 15.6\ \text{dB}$$

Appendix B. MRC Multi-centre Otitis Media Study Group

The group does not have a formal constitution; regular meetings of the core staff and collaborating consultants were only held during the planning and conduct of the trial in the mid-late 1990s. This list satisfies editorial requirements for authorship and governance and human resource requirements for acknowledged contributorship.

B.1. MRC scientific staff

Project Leader: Haggard MP (Guarantor of present paper and co-author); Health Services Researcher/Trial Co-ordinator: Gannon MM; Co-ordinator: Birkin JA; Statisticians: Bennett KE, Nicholls EE, Spencer H (Statistician for analyses used in this paper); Otolaryngology research fellows: Georgalas C, Daniel M; Audiological Scientist: Higson JM (Co-author); Psychologists: Smith SC, Hind SE; Epidemiologist: Rovers MM.

B.2. Academic medical staff

Lead Academic Clinician: Browning GG; Attached otolaryngology research fellows: Georgalas C, Daniel M.

B.3. MRC support staff

Data Manager: Egner EM; Research Assistants: Hayman T, Greenwood DC, Carroll RA, Jones H, Richmond TB, Wade AR, Braham L, Moorjani P, Pearson DAS, Kirk G; Audiologist: Baskill JL.

B.4. RCT centres

Royal Victoria Hospital, Belfast; Ulster Hospital, Dundonald; University Hospital, Nottingham; Leicester Royal Infirmary, Leicester; Royal Hospital for Sick Children, Bristol; Freeman Hospital, Newcastle; Royal Hospital for Sick Children, Edinburgh; Queen Alexandra Hospital, Portsmouth; Sheffield Children's Hospital, Sheffield; Coventry and Warwickshire Hospital, Coventry; University Hospital of Wales, Cardiff.

B.5. Other contributing centres

Royal Hospital for Sick Children, Glasgow; Manchester Children's Hospitals; Diana, Princess of Wales and Heartlands Hospitals, Birmingham; Epsom General Hospital, Epsom; Sunderland Royal Hospital, Sunderland; Tyrone County Hospital, Omagh.

B.6. Collaborating consultants

Adams DA; Toner JG; Gibbin KP, Jones NS, Birchall JP; Thomas RSA, Narula AA, Murty GE; Griffiths MV; Wilson JA, Meikle D, Johnson IJM, Date S; Kerr AIG; Robinson D, Madden G; Bull PD, Chapman DF, Parker AJ; Curry AR; Williams RG; Morrissey MSC, Geddes NK; Zarod AP, Willatt DJ; Proops D, Pearman K; Robb PJ; MacKee I; Law K.

B.7. Nurses/Research assistants

Casey M; Megarry A; Smith L, Smethurst H; Pearce G, Hook P, Phipps J; Gibbs S, Dunster S; Brown E, Birch EM, Page V; Sim A; Sharples G, Crawford J; Davies C; Bartolo J; Edwards C; Wilson J, Johnson N; Baxendale C, Welsby W, Pickavance N; Cox V, Markham K; Robb JA.

B.8. Audiological scientists/audiologists

Flanagan R, Housten G, Jordan G, Bingham M, Cassidy J; Burns C, Hamill C, Glover L, Eccles L; Armstrong-Bednall G, Towle B, Dyer K, Carr J, Herrod J, Levick A; Dunmore P, Parker-Emery C, Sutton R, Gordon C; Midgley L, Barnett N, Foley R, Machlin R; Elliott C, Davidson T, Walker L, Keith J; Brady M, Barr-Hamilton R, Wilson S, McKenzie J, Edgar R, Noble H; Thomas H, Paul B, Gilbert M, Hoblyn R, Walter S, Griffiths E, Inge F; Hodges M, Lawrence C, Walsh A, Loxley T, Byrom P; Walker D, Najaran R, Blyth G, Mace L; Rabaiotti W; Trearty C, Carmichael A; Ronson R, France K, Pilkington H; Hilton F, Parsons J; Whitby G, Robinson S; Brown E, James P; Allen R.

References

[1] J. Bamford, K. Uus, A. Davis, Screening for hearing loss in childhood: issues, evidence and current approaches in the UK, J. Med. Screen. 12 (2005) 119—124.
[2] H.M. Fortnum, A.Q. Summerfield, D.H. Marshall, A.C. Davis, J.M. Bamford, Prevalence of permanent childhood hearing impairment in the United Kingdom and implications for universal neonatal hearing screening: questionnaire based ascertainment study, BMJ 323 (2001) 536—539.
[3] M.P. Haggard, B. McCormick, M.M. Gannon, H. Spencer, The pediatric otologic caseload resulting from improved screening in the 1st year of line, Clin. Otolaryngol. 17 (1992) 34—43.
[4] Y.K. Kemaloğlu, B. Grűndűz, S. Gökmen, M. Yilmaz, Pure tone audiometry in children, Int. J. Pediatr. Otorhinolaryngol. 69 (2005) 209—214.
[5] M. Haggard, J. Higson, H. Spencer, Bone-conduction measurements in children with OME—an issue for professional debate, BSA News (July) (2007) 33—37.
[6] A. Noone, M. Goldacre, A. Coulter, V. Seagraott, Do referral rates vary widely between practices and does supply of services affect demand? A study in Milton Keynes and the

Oxford Region, J. Roy. College Gen. Pract. 39 (1989) 404—407.
[7] M. Ryan, G. Mooney, Research in medical practice variations: where now? A paper for debate. Health Economics Research Unit, University of Aberdeen, discussion paper no. 04 (1991).
[8] MRC Multi-Centre Otitis Media Study Group, Sensitivity, specificity and predictive value of tympanometry in predicting a hearing impairment in otitis media with effusion, Clin. Otolaryngol. Allied Sci. 24 (1999) 294—300.
[9] MRC Multi-centre Otitis Media Study Group. Influence of age, type of audiometry and child's concentration on hearing thresholds. Br. J. Audiol. 34 (2000) 231—240.
[10] MRC Multi-centre Otitis Media Study Group. Surgery for persistent otitis media with effusion: generalizability of results from the UK trial (TARGET), Trial of alternative regimens in glue ear treatment, Clin. Otolaryngol. 26 (2000) 417—424.
[11] MRC Multi-centre Otitis Media Study Group. Risk factors for persistence of bilateral otitis media with effusion. Clin. Otolaryngol. Allied Sci. 26 (2000) 147—156.
[12] MRC Multi-centre Otitis Media Study Group, Air Conduction Estimated from Tympanometry (ACET) 1: Relationship to measured hearing, Int. J. Pediatr. Otorhinolaryngol., 2008.
[13] BSA, Recommended procedures for pure-tone audiometry using a manually operated instrument, Br. J. Audiol. 15 (1981) 213—216.
[14] British Society of Audiology, Recommended Procedure: Pure Tone Air and Bone Conduction Threshold Audiometry With and Without Masking and Determination of Uncomfortable Loudness Levels, The British Society of Audiology http://www.thebsa.org.uk/docs/bsapta.doc (accessed on March 2004).
[15] J. Jerger, Clinical experience with impedance audiometry, Arch. Otolaryngol. Head Neck Surg. 92 (1970) 311.
[16] G.A. Zielhuis, G.H. Rach, P. Van den Broek, The occurrence of otitis media with effusion in Dutch preschool children, Clin. Otolaryngol. 15 (1990) 147—153.
[17] A.A. Timmerman, C.M. Meesters, R. Speyer, L.J. Anteunis, Psychometric qualities of questionnaires for the assessment of otitis media impact, Clin. Otolaryngol. 32 (2007) 429—439.
[18] A.M. Tharpe, F.H. Bess, Minimal, progressive, and fluctuating hearing losses in children, characteristics, identification, and management, Pediatr. Clin. North. Am. 46 (1999) 65—78.
[19] F.H. Bess, A.M. Tharpe, A.M. Gibler, Auditory performance of children with unilateral hearing loss, Ear Hear. 7 (1986) 20—26.
[20] S. Gatehouse, M.P. Haggard, The effects of air—bone gap and presentation level on word identification, Ear Hear. 8 (1987) 140—146.
[21] R.J. Ruben, Host susceptibility to sequelae, in: R.M. Rosenfeld, C.D. Bluestone (Eds.), Evidence-based Otitis Media, BC Decker, Hamilton, 2003, pp. 505—514.
[22] MRC Multi-centre Otitis Media Study Group, The role of ventilation tube status in the hearing levels in children managed fro bilateral persistent otitis media with effusion, Clin. Otolaryngol. Allied Sci. 28 (2003) 146—153.