Optimal Outcome Measures, Research Priorities, and International Cooperation

Robyn Cox, Martyn Hyde, Stuart Gatehouse, William Noble, Harvey Dillon, Ruth Bentler, Dafydd Stephens, Stig Arlinger, Lucille Beck, Deborah Wilkerson, Sophia Kramer, Patricia Kricos, Jean-Pierre Gagné, Fred Bess, and Lillemor Hallberg

The participants in the Eriksholm Workshop on "Measuring Outcomes in Audiological Rehabilitation Using Hearing Aids" debated three issues that are reported in this article. First, it was agreed that the characteristics of an optimal outcome measure vary as a function of the purpose of the measurement. Potential characteristics of outcome self-report tools for four common goals of outcome measurement are briefly presented to illustrate this point. Second, 10 important research priorities in outcome measurement were identified and ranked. They are presented with brief discussion of the top five. Third, the concept of generating a brief universally applicable outcome measure was endorsed. This brief data set is intended to supplement existing outcome measures and to promote data combination and comparison across different social, cultural, and health-care delivery systems. A set of seven core items is proposed for further study.

(Ear & Hearing 2000;21;106S–115S)

As the other papers in this supplement attest, the Eriksholm Workshop on Measuring Outcomes in Audiological Rehabilitation Using Hearing Aids focused on using and promoting outcome measurement and designing and selecting appropriate tools. In addition to the deliberations set forth in the presentations, the participants undertook to debate the following three questions:

1. How do the optimal characteristics of outcome measures vary as a function of measurement goals?
2. What are the most urgent research needs in outcome measurement for audiological rehabilitation?
3. Should there be a small set of core outcome items that could supplement existing measures to promote data comparison across different contexts?

This article reports the results of these discussions.

Optimizing Outcome Measures

To what extent do the optimal characteristics of outcome measurement vary depending on the purpose of the measurement? To explore this matter, workshop participants deliberated the desirable qualities of an outcome measurement tool for four different but common applications of outcome data. Two to four participants decided how to meet each goal and then presented their plan to the group. Not surprisingly, there was not unanimous agreement on the most preferred characteristics of any type of outcome measure. Nevertheless, it was possible to gain reasonable consensus approval for most of the principles reported here. On occasion, there were strong but irreconcilable views held by participants. These are reported under "Further considerations and other opinions."

Goal 1: To Assess the Rehabilitative Outcomes for an Individual Hearing-Impaired Person

This application focuses on determining outcomes for a particular client within a clinical setting. For the purposes of this discussion, it was decided to define a measure that would be applicable to adult, noninstitutionalized hearing aid wearers who have hearing loss within the mild to severe range. It is intended to be suitable for use with the vast majority of elderly first-time hearing aid wearers and, with some adjustments in interpretation, with experienced wearers.

University of Memphis (R.C.), School of Audiology and Speech-Language Pathology, Memphis, Tennessee; Mt. Sinai Hospital (M.H.), Toronto, Ontario, Canada; MRC Institute of Hearing Research (S.G.), Glasgow, Scotland; School of Psychology (W.N.), University of New England, Armidale, NSW, Australia; National Acoustical Laboratories (H.D.), NSW, Australia; University of Iowa (R.B.), Speech Pathology and Audiology, Iowa City, Iowa; Welsh Hearing Institute (D.S.), University Hospital of Wales, Cardiff, Wales, United Kingdom; University Hospital (S.A.), Department of Audiology, Linköping, Sweden; VA Medical Center (L.B.), Washington, DC; CARF (D.W.), The Rehabilitation Accreditation Commission, Tucson, Arizona; University Hospital VU (S.K.), Department of Audiology, Amsterdam, The Netherlands; Communication Sciences & Disorders (P.K.), University of Florida, Gainesville, Florida; Université de Montréal (J.-P.G.), Ecole d'orthophonie et d'audiolgie, Montréal, Canada; Vanderbilt University Division of Hearing and Speech Sciences (F.B.), Nashville, Tennessee; and The Nordic School of Public Health (L.H.), Göteborg, Sweden.

0196/0202/00/2104-0106S/0 • Ear & Hearing • Copyright © 2000 by Lippincott Williams & Wilkins • Printed in the U.S.A.

General Recommendations • The measure should be designed to facilitate assessing the client's needs and preferences; planning the rehabilitation treatment; deciding when the rehabilitation program should be concluded; and evaluating the success of the program, in terms of the daily life consequences for the client and family.

Because the measure should enhance, as well as simply assess, the quality of the rehabilitative outcome, it can be valuable for this type of outcome measure to be administered and interpreted by the clinician providing the rehabilitative services. In this way, information obtained during the measurement process can be used to improve planning and treatment.

The measure should primarily be based on individualized self-report, in which the nature of the assessment items or listening situations is determined by each client, rather than being standardized across clients. The optimal measure will assess benefit in terms of both disability (activity limitation) and handicap (participation restriction) reduction and will also include an assessment of satisfaction. In addition, considerable importance is placed on the application of a parallel measure designed to determine the outcomes of the rehabilitation program from the point of view of others in the client's life. This is addressed by soliciting self-report from a significant other person (SO) accompanying the client.

Illustration/Example • This type of measure could be accomplished using a brief two-part inventory composed of 1) a version of the Client Oriented Scale of Improvement (COSI) (Dillon & Ginis, 1997) modified to specifically address issues of disability (activity limitation) and handicap (participation restriction), and 2) an overall item to assess satisfaction.

For example, at the initial interview, the client might be asked to:

"Tell me about the problems your hearing is causing you that you would like to solve."

Experience with the COSI indicates that this question predominantly elicits reports of disabilities. The clinician's task is to write down the problems mentioned by the client. If the client mentions general situations, such "I have trouble hearing people in noisy places," the clinician should ask for a specific example that is particularly important to the client. After each item is recorded, the clinician should ask whether there are any other problems that the client would like to solve. Three to five items should preferably be identified, although the method can be used with a fewer or greater number of items. Decreased accuracy of the overall assessment is likely with fewer items.

Following identification of disabilities, the issue of handicaps might be approached with a question such as:

"Have these problems with your hearing caused you to change the things you do or the way you do them?"

If possible, this should be explored for the identified disabilities and in general. Also, the clinician should be aware that some handicapping conditions such as psychological effects are not directly related to activity limitation and, so, must be explored independently of disabilities. That is, the client might be queried as to levels and kinds of emotional distress experienced as a result of their hearing problems.

Next, if the client is accompanied by an SO, the clinician might ask the SO:

"Hearing loss often causes problems for family members as well. Can you tell me about any problems that you want to solve from your point of view?"

The SO's responses should also be refined, if necessary, so that specific examples from daily life are elicited.

In addition to providing the framework for outcome assessment at the conclusion of the rehabilitation program, this information obtained at the pretreatment interview can be used to assess the apparent motivation of the client to obtain rehabilitation; it can also serve to check any divergence between the level of the measured hearing impairment (i.e., the audiogram) and the level of disability implied by the stated needs of the client; as well as any divergence between the level of disability or handicap stated by the client relative to that stated by the SO. These issues are helpful in designing a rehabilitation program.

Outcome assessment occurs first at the final appointment, when the clinician tentatively believes the rehabilitation program should be concluded. Outcome assessment of longer-term changes should rely on follow-up at some time after the conclusion of services (e.g., 3 to 6 mo later). For each of the previously identified disabilities and handicaps, and at each data collection point, the client is asked the following questions:

"When we first talked about your hearing problems, you said: (refer to initial interview). Can you tell me how much change there has been for you in this situation?" (or "in this regard," as appropriate to the nature of the item).

The client is asked to classify the change as one of Worse, No difference, Slightly better, Better, or Much better. Responses mid-way between these categories are acceptable if the client is wavering between two adjacent categories.

The SO is asked the same questions, appropri-

ately modified to address the concerns that he/she had expressed at the initial interview.

Finally, satisfaction should be assessed using a global question, such as: "Overall, how satisfied are you with your new hearing aids?" The list of responses includes: Very Satisfied, Satisfied, Neutral (neither satisfied nor dissatisfied), Dissatisfied, Very Dissatisfied. Experience with this item suggests that clients aggregate across their perceptions of quality and value to respond to this item. If the client gives a low score, the clinician should ask which aspect(s) is not satisfactory. The SO should also be asked about his/her satisfaction with the results of the rehabilitation program, from his/her point of view.

Further Considerations and Other Opinions • This approach to outcome measurement does not require that data obtained from different clients be cumulated or compared. The focus here is on planning for the individual, though as we will see in the next section, individual-level data can be used to describe clinical effectiveness. Quantification and cumulation of personalized data is possible, as described by Dillon, Birtles, and Lovegrove (1999). Although such quantification does not separate better-than-average from average performance very successfully, it is quite well suited to identifying individuals who are receiving less-than-average improvements.

We have suggested for pragmatic reasons that this type of outcome data be collected by the clinician that is responsible for providing rehabilitative services. However, there is the potential that the client's responses will be influenced by the lack of anonymity in this procedure. For this reason, some professionals might argue that the worth of an intervention should be assessed independently of the service provider to encourage candid feedback from the client. This approach would require an entirely different type of outcome measure from the one described here and it would have the disadvantage that the data could not be used to promote the rehabilitative outcome for this particular client.

Goal 2: To Assess the Effectiveness of the Services Provided by a Particular Clinical Unit or Agency

This type of measure may be used by clinics for audit purposes and also by service purchasers to evaluate the effectiveness of particular service providers. The interventions assessed would include all aspects of audiological rehabilitation including hearing aids and cochlear implants as well as noninstrumental rehabilitation and assistive devices (environmental aids). What is being assessed is the effectiveness of the intervention rather than its efficiency (cost-effectiveness). The latter can be derived from the effectiveness assessed in terms of all the direct and indirect costs involved. This normally should be conducted by an auditor or accountant on a regular if infrequent basis. Alternatively, these data can be derived from a well-planned administrative data set.

General Recommendations • The assessment should be aimed at evaluating the rehabilitation service provision of the agency overall rather than a specific clinician or client. Two alternatives are available to administrators wishing to generate agency-wide outcome data. First, they can distribute a standardized questionnaire in the mail. In this case, the time frame for data collection should be controlled. An interval of 3 to 6 mo after the rehabilitation program was completed seems reasonable. Second, they can request each clinician to obtain outcome data on all or a random sample of his/her own clients. In this case, simplicity of administration together with the premise that the test(s) do not add markedly to the session time, is essential to ensure that the outcome measure is appropriately administered by all services and not just by academic departments and those with a research interest.

Whichever alternative is selected, the key to success of such a measure lies in its amenability to quantification and accumulation of data across clients.

Illustration/Example • If it is determined that each clinician will secure outcome data on all clients and these will then be accumulated, a personalized self-report scale based on the COSI (Dillon & Ginis, 1997) such as described under Goal 1 could be used. This would have the advantage that it is clearly relevant to the patient's needs and takes little time to complete over and above activities routine for the treatment program. Dillon et al. (1999) further indicated that it is broadly acceptable to a wide range of hearing health-care professionals, and that it related well to a pool of outcome measures in the Australian population (this comparison does, however, need to be repeated in other countries).

Use of a personalized outcome measure has the disadvantage of being less clearly quantifiable. However, suggestions for quantifying the COSI have been provided by Dillon et al. (1999) and could be adapted for this use. The outcome measure should include at least three and preferably five areas, defined at the beginning of the intervention, for which the client is seeking help. Each is then assessed at the last session both in terms of degree of change (on a 5-point scale) and in terms of final ability (again on a 5-point scale). The mean result for each topic is then calculated for both change and final ability.

At the agency level, satisfaction assessment

should comprise at least two items. The first should tap the individual's satisfaction with the way that they have been treated by the service—for example:

"How satisfied were you with the approach of the professionals you saw about your hearing problems?"

The second should tap their satisfaction with the outcome of the process—for example:

"How satisfied are you with the extent to which your hearing problems have been reduced?"

Both items should be scored using a five point scale from Very Satisfied to Very Dissatisfied (see Goal 1).

The final part of the in-clinic outcome assessment would ask the individual to list the benefits and shortcomings of the rehabilitative process that they had undergone, as used by Stephens and Meredith (1991) in a hearing aid study and later in a range of other studies. The wording would be:

"Please make a list of any benefits which you have obtained from the audiological rehabilitation you have received. Write down as many as you can think of" and *"Please make a list of any shortcomings of the audiological rehabilitation that you received. Write down as many as you can think of."*

Common themes across respondents, both positive and negative, would reveal areas in which service provision is regarded as working well or poorly. Common themes would also show strengths of different providers relative to each other.

Further Considerations and Other Opinions • The illustration used here relies on collection of outcome data in the clinical context. As noted above, Goal 2 can also be addressed using post-treatment mailings. In this case it would be necessary to make use of a standardized inventory to secure the data.

Any assessments comparing different hearing services will need to take into account the case mix seen in the particular department. The case-mix will cover factors including the age range, severity and duration of hearing impairment, gender and ethnic balance, social class, education level, and whether a first-time or returning client.

In addition, records should note the type of referrals seen whether primary, secondary or tertiary referrals. These entail, respectively, self-referral, referral from a primary physician (general practitioner), or referral from another Audiology department for specialist assessment and/or intervention. This information will be helpful in targeting future marketing efforts.

Goal 3: To Assess the Effectiveness of New Hearing Aid Technologies

This outcome measurement task must identify characteristics of outcome self-assessment tools that would be relevant in the context of a clinical trial of a novel hearing aid technology. It is assumed that small-scale laboratory studies have demonstrated the efficacy of the new "treatment." The purpose of the clinical trial is to determine whether the laboratory findings are sustained under real-world conditions. This involves a new study using a larger number of subjects who are representative of the intended target population who might use the new strategy. It also involves device use in field conditions with diverse acoustical characteristics, as well as the use of subjective rating scales as outcome measures. The null hypothesis is that there is no treatment effect, namely that the new technology is not superior to the comparison condition. The experimental hypothesis is that the new hearing aid is superior by at least a minimum amount, the minimum treatment effect size.

General Recommendations • Because the emphasis in this sort of goal is assessment of the effectiveness of a technical device, a standardized rather than client-oriented form of inquiry is required. This enables a test of the device with particular reference to the conditions in which the device is claimed to offer advantages. The objective is to be able to detect the minimum treatment effect when it is genuinely present with a probability (statistical power) typically set to at least 0.80. In selecting a self-report outcome measure, the simplest course is to try to identify an existing standardized device that is suitable to the task. To achieve acceptable levels of validity, reliability and responsiveness (see Hyde, 2000), the following attributes are likely to be necessary.

Content Domain Relevancy: The self-report scale should include a subscale that targets the key issue postulated to differentiate the treatment conditions, for example, understanding speech in noise. The more relevant the subscale's content in relation to the question motivating the trial, the better the reliability and sensitivity to treatment effects are likely to be.

Appropriate Difficulty: The scale also should be appropriate for the trial context and population in the level of listening difficulty represented in the test items The key point is to aim for realism in the situations represented by test items. Another aspect of difficulty is the reading level required to interpret properly the instructions and the item content. Reading level should not be more demanding than that needed to understand popular newspapers or magazines.

Item Applicability: For useful measurement in this type of trial context, it is important that each subject perceives all or most of the test items as both relevant to the underlying problem and applicable

personally (face validity). The item non-response frequency should be low.

Secondary Coverage: It is desirable that the outcome measurement device addresses important secondary areas that could influence the further development and application of the primary treatment. Possible examples are issues such as loudness discomfort and sound quality.

Illustration/Example • Assume that the new technology is postulated to improve ability to understand speech in noisy environments. An example of a self-assessment device that seems well-suited for the clinical trial, based on the considerations outlined above, is the Shortened Hearing Aid Performance Inventory for the Elderly, or SHAPIE (Dillon, 1994; Walden, Demorest, & Hepler, 1984). Although not the only scale that might be considered, the SHAPIE possesses many of the desirable attributes noted earlier. The development and validation work for the SHAPIE focused on item refinement and selection to increase content relevance for elderly subjects, to achieve a broad range of scores over subjects, high internal consistency, simplicity of factor structure and high test-retest reliability.

The SHAPIE exists in 25-item and 40-item forms, each item using a 5-point ordinal scale to rate aid helpfulness in a variety of situations. It focuses on hearing speech in general, with a subscale on hearing in competing conditions. For the 25-item form, the 95% Critical Difference between two administrations in the individual subject is 0.55 scale points when subjects are preselected to include only those for whom most of the SHAPIE items are relevant. A difference outside this range is interpreted as evidence of a real difference in scores. The number of subjects needed can be determined based on the published data for the SHAPIE. If we assume that the smallest interesting effect size is 0.2 scale intervals, 17 subjects per group will be needed to provide a power of 0.8 for the clinical trial. To further optimize the design of the trial, one could select subjects whose major problem was hearing speech in competing noise. However, this would compromise, to some extent, the generalizability of the trial results.

Further Considerations and Other Opinions • This illustration has been deliberately simplified to emphasize the issues that must be considered when it is desired to collect data to support a particular claim. In real world applications of these concepts it is important to keep in mind that there might be other dimensions or domains within which the benefits or detriments of the new technology might be evident. For example, improving ability to understand speech in noisy environments can have manifestations (including demerits) in speech identification abilities, ease of listening, sound quality, and spatiality effects. In other words, there may be advantages and disadvantages from a particular hearing aid processing strategy that might lie in unexpected and unpredictable domains. This is a strong argument in favor of designing clinical trials to be as comprehensive as possible in terms of the dimensions assessed.

Goal 4: To Evaluate the Effectiveness of Hearing Rehabilitation Services on Quality of Life

A quality of life measure (more strictly speaking, a health-related quality of life measure—elsewhere referred to as functional health status) is, by definition, generic in nature and refers to no particular disease or condition, or no particular intervention for that disease or condition. Thus, this application of self-report outcome measurement will often be used to compare the effectiveness and cost-effectiveness of auditory rehabilitation, including amplification provision, with other interventions for other conditions.

The purposes for which one might wish to compare effectiveness and cost-effectiveness across rather than within conditions and interventions reside at the policy level where resource-allocation decisions are made. These policy decisions are not taken at frequent intervals but, rather, are made by appropriate funding agencies when reviews of resource allocation are made. Thus, the information would not necessarily be gathered on a routine basis on all patients or clients processed by a particular facility. Therefore, the measures do not necessarily have to be compatible with routine clinical practice on all patients or clients.

General Recommendations • A formal cost-utility analysis requires a "health utility" scale (Torrance, 1986; Torrance & Feeny, 1989), and integration and discounting across time allows us to derive measures of Quality Adjusted Life Years. Although there are many measures that purport to access health-related quality of life or functional health status (Bess, 2000), few are demonstrably appropriate for this particular purpose. Probably the two preeminent ones are the Short Form Health Survey SF36 (Brazier, Usherwood, Harper, & Thomas, 1998; Ware & Sherbourne, 1992) and the Health Utilities Index (Torrance, 1986). However, each of these measures has limitations for our particular application.

Most scales of health-related quality of life were constructed, either with severe life-threatening disease in mind, or to load relatively heavily on aspects of physical and motor function. This is particularly

true of the SF36, which contains very few items with the potential to be influenced by the psychological, social, and emotional consequences of sensory deficits.

The Health Utilities Index does load to some extent on the psychosocial consequences of auditory deficits. However, there are significant concerns about the accuracy of the weightings used to transform questionnaire responses to overall utility scales. These are derived from assessments of various health states made by individuals who may not have actually experienced the health states that they are asked to evaluate. This is particularly problematic in the hearing disability and handicap domain, because the consequences of communication deficits are largely unrecognized by individuals who have not experienced them. For example, it is commonly accepted by normally hearing and normally sighted individuals that visual deficits are "more damaging" than auditory deficits. This can be illustrated using an example from the Health Utilities Index. The utility of a respondent's health state is derived by mapping responses on the questionnaire onto particular levels on eight dimensions. An algorithm then converts the value into a utility. These mapping rules combine limitations in vision, hearing and speech production into a single value of a "sensation" attribute. The rules place a respondent at Level 4 on the sensation attribute if they are at Level 6 of the vision attribute, irrespective of the degree of hearing disability. This scoring method implies that vision problems are much more damaging than hearing problems to quality of life. By this reasoning, being totally deaf and totally blind is no worse than being totally blind and having perfect hearing. This result is counter-intuitive.

The low occurrence of items that have the potential to be affected by communication deficits, combined with the under-estimated impact of such deficits on health status, leads to a deficiency in the overall appropriateness of both the SF36 and the Health Utilities Index for assessing the effectiveness of audiological rehabilitation. This is particularly true of the SF36, and may well account for some of the early negative findings cited by Bess (2000). Research is urgently needed in collaboration with the Health Services Research and Health Economic communities to address the issues associated with both item content and item weighting. Such concerns affect not only the audiological community, but also other groups within health-care provision whose conditions and interventions are inappropriately represented within current utility scales. This research priority requires collaboration across disciplines if it is going to carry appropriate weight and achieve acceptance by the policy-making processes within the different health-care systems.

Illustration/Example • An appropriate target population and setting includes all adults whose management of hearing disability and handicap involves provision of personal amplification and other auditory rehabilitation and who are physically and mentally capable of self-completion of a questionnaire.

Research collaboration is critical to develop a measure to determine the utility and cost-utility of audiological interventions. In the meantime, we suggest that the SF-36 is not appropriate for that purpose. The Health Utilities Index, despite its drawbacks in construction and validation, would be the current instrument of choice, given its established use and acceptance in the associated arena of cochlear implantation (Summerfield & Marshall, 1995).

Further Considerations • Those concerned with selecting outcome measures should note that it would be quite inappropriate to select a generic health-related quality of life measure to compare the clinical effectiveness of different auditory rehabilitation options. This is due to the inevitable sacrifice of sensitivity that is associated with generic rather than condition-specific instruments. Comparison of different auditory rehabilitation regimes would be most appropriately served by condition-specific measures that specifically tap into the domains of auditory disability (activity limitation) and handicap (participation restriction).

RESEARCH NEEDS IN OUTCOME MEASUREMENT

At the conclusion of the Workshop, a list of research needs was generated and then prioritized using a rating method. The top 10 research priorities are presented below, with a brief discussion of the first five.

1. Explore the relationship between expectations and outcome, especially including satisfaction. The relationship between prefitting expectations and post-treatment outcomes of hearing aid fitting intrigues clinicians because of the potential implications for treatment decisions. Some investigations exploring this relationship have been reported (e.g., Brooks & Hallam, 1998; Cox & Alexander, 2000; Schum, 1999) but much more remains to be done. Further, we should keep in mind that cultural differences and differing health-care systems in different countries might affect the relationship. Thus, the relationship between expectations and motivations in the

UK National Health Service might be different from that in a private-led environment such as the United States.

2. Determine the relationship of COSI to other outcome measures in multiple countries. In the paper by Dillon and Ginis (1997), in which COSI was first described, a strong relationship was found between COSI and an aggregate of other outcome measures. COSI has two further advantages: It is client oriented, and it is brief. These qualities have been seen to encourage compliance in its completion by a wide range of audiological departments (Dillon & So, 2000). They also encourage clinicians to think in terms of what is relevant for the specific client. However, because it is a nonstandardized measure, its face validity, together with its generality in comparing different departments, may be questioned by some in Public Health Medicine in other countries. To explore this matter, there is a need for further studies within other centers, to compare COSI outcomes with other widely used and perhaps standardized measures. This will help to determine whether the conclusions from Australia can be generalized to other sociomedical systems.
3. Delineate the effects of extra-audiological factors in outcome measurement. As reviewed by Kricos (2000), some evidence about the effects of personality, gender, etc., on outcomes has been reported. Nevertheless, a clear and parsimonious statement about the extra-audiological determinants of outcome and their relative contributions to the effectiveness of treatments has yet to be offered by the research community in a package that can be utilized at the clinical level. Knowing what these factors are is also crucial to risk-adjustment of outcome findings for performance indicators and other quality comparisons.
4. Determine the generic components of quality of life that are affected by a hearing problem. Few research data exist to delineate the ways in which, and dimensions on which, hearing impairment impinges on quality of life of individuals and quality of life within families and SOs. Though varying with individual circumstances, hearing loss would be expected to have its greatest impact on the quality of family life, the quality of the social world beyond the family (friends, acquaintances, others), the quality of work and/or community life, and various dimensions of personal efficacy. Given the competition among services for limited health-care resources, it is imperative for our field to develop a more detailed and sensitive understanding of these issues. This type of data will not only facilitate more effective rehabilitation strategies, but also promote development of more accurate quantification of utilities of auditory deficits.
5. Develop a client-oriented instrument that specifically evaluates both disability and handicap. In patient-generated questionnaires such as the problem questionnaire (Barcham & Stephens, 1980), and COSI (Dillon & Ginis, 1997), complaints relating to Participation Restriction (handicap) are relatively rarely listed compared with those related to Activity Limitation (disability) (Dillon et al., 1999; Stephens, Jones, & Gianopolus, 2000). Nevertheless, there is some evidence to suggest that it is questions of participation restriction that encourage an individual to seek help at a clinic (to a greater extent than activity limitations). It can thus be argued that to obtain a fully appropriate client-centered outcome measure, efforts should be made to use a wording that will specifically elicit a Participation Restriction element. The goals of the research would be: to devise a suitable wording, to quantify the degree to which it elicits rehabilitation needs related to Participation Restriction, and to evaluate the validity of the new measure relative to existing measures. Some preliminary ideas regarding suitable wording were outlined within Goal 1 of this paper.
6. Study comorbidity in the elderly hearing impaired and impact on condition-specific handicap (participation restriction) and quality of life.
7. Study postintervention long-term time course of disability (activity limitation) and handicap (participation restriction).
8. Gather utility values in hearing impairment using standard-gamble and time-trade techniques for a variety of populations.
9. Derive a system of classifying different types of hearing aid benefit and indications for measurement of each type.
10. Explore the relationships among functional outcomes and satisfaction (with outcome, process, care and overall).

A "UNIVERSAL" OUTCOME MEASURE

Although many investigations collect outcome data, it is often difficult to pool data across studies because of protocol and instrument differences. Fur-

ther, as pointed out by Arlinger (2000), outcome questionnaires may be sensitive to a variety of factors related to the cultural and social environment, making international comparisons suspect. Thus, the potential increases in power that can be derived by aggregating data across time and place cannot often be realized. This is especially unfortunate in our field because audiological rehabilitation research is frequently limited to small subject groups, and self-report outcome measures are not highly sensitive instruments. Thus, it is often difficult to arrive at unequivocal answers to research questions due to limits of statistical power. Recognizing this limitation, workshop participants endorsed, in principle, the goal of devising a minimal set of core outcome items that would be sufficiently general to apply to many different types of investigations carried out in different countries in the world. We have optimistically called these the International Outcome Items for Hearing Aids (IOI-HA). This core set of items is not intended to replace the outcome measure(s) selected as optimal for any particular study. Instead, it is intended to function as a useful addendum to existing measures in a research context. There might also be potential for the set to function as a standalone tool for quality assessment. We encourage the addition of these core items to the set of outcome measures utilized in any investigation of audiological rehabilitation using hearing aids. In the future, data on these items from various studies can be pooled in meta-analyses or compared across health-care models, countries, processing types, etc.

Psychometric techniques used in selecting and refining the items included: cueing the question intent, where possible, in the response options (even at the cost of slight awkwardness); identifying a concrete time window for reflection (if the time frame of 2 wk is too short for a given study, the time window might need to be modified in that case); wording the question as directly as possible in relation to the intended construct; and minimizing reversal.

Efforts were made to avoid substantial duplication of items from popular instruments, and the number of items was pared down so that the burden of including them in a study design would minimized. An important and continuing concern is the level of intellectual and verbal sophistication required to respond to the items. In North America, a Grade 4 reading level is a common design maximum, and this set goes beyond that, despite efforts of the contributors. Another major concern was the need to formulate the questions in a way that would make them equally applicable in any country and cultural-social environment where they might be used. Future developments of the IOI-HA will include translation in to other languages. Efforts are already underway to translate the items into Dutch, Spanish and Swedish. However, the translation process will not be trivial and it remains to be seen whether translations can be accomplished that retain the same nuances as the English items.

The items of the IOI-HA appear in the Appendix. They are practically oriented. They comprise a mini-profile more than a scale. The psychometric characteristics of the set are unknown at this time but several workshop participants will contribute to data on this topic. It is hoped that by calibration on well-defined test populations across the range of hearing loss and age in different countries, the IOI-HA will eventually prove to be a valuable addition in this field.

Address for correspondence: Robyn Cox, Ph.D., Memphis Speech & Hearing Center, 807 Jefferson Avenue, Memphis, TN 38105.

Received March 30, 2000; accepted March 30, 2000

REFERENCES

Arlinger, S. (2000). Can we establish internationally equivalent outcome measures in audiological rehabilitation? *Ear and Hearing, 21*(Suppl).

Barcham, L. J., & Stephens, S. D. G. (1980). The use of an open-ended problems questionnaire in auditory rehabilitation. *British Journal of Audiology, 14*, 49–54.

Bess, F. (2000). The role of generic health-related quality of life measures in establishing audiological rehabilitation outcomes. *Ear and Hearing, 21*(Suppl.).

Brazier, J., Usherwood, T., Harper, R., & Thomas, K. (1998). Deriving a preference-based single index from the UK SF-36 Health Survey. *J Clinical Epidemiology, 51*, 1115–1128.

Brooks, D. N., & Hallam, R. S. (1998). Attitudes to hearing difficulty and hearing aids and the outcome of audiological rehabilitation. *British Journal of Audiology, 32*, 217–266.

Cox, R. M., & Alexander, G. C. (2000). Expectations about hearing aids and their relationship to fitting outcome. *Journal of the American Academy of Audiology, 11*, in press.

Dillon, H. (1994). Shortened Hearing Aid Performance Inventory for the Elderly (SHAPIE): A statistical approach. *Australian Journal of Audiology, 16*, 37–48.

Dillon, H., Birtles, G., & Lovegrove, R. (1999). Measuring the outcomes of a national rehabilitation program: Normative data for the Client Oriented Scale of Improvement (COSI) and the Hearing Aid User's Questionnaire (HAUQ). *Journal of the American Academy of Audiology, 10*, 67–79.

Dillon, H., & Ginis, J. A. (1997). The Client-Oriented Scale of Improvement (COSI) and its relationship to several other measures of benefit and satisfaction provided by hearing aids. *Journal of the American Academy of Audiology, 8*, 27–43.

Dillon, H., & So, M. (2000). Incentives and obstacles to the routine use of outcomes measures by clinicians. *Ear and Hearing, 21*(Suppl.).

Hyde, M. (2000). Reasonable psychometric standards for outcomes self-report scales. *Ear and Hearing, 21*(Suppl.).

Kricos, P. B. (2000). The influence of nonaudiological variables on audiological rehabilitation outcomes. *Ear and Hearing, 21*(Suppl.).

Schum, D. J. (1999). Perceived hearing aid benefit in relation to perceived needs. *Journal of the American Academy of Audiology, 10*, 40–45.

Stephens, D., Jones, G., & Gianopolus, I. (2000). The use of outcome measures to formulate intervention strategies. *Ear and Hearing, 21*(Suppl.).

Stephens, S. D. G., & Meredith, R. (1991). Qualitative reports of hearing aid benefit. *Clinical Rehabilitation, 5*, 225–229.

Summerfield, A. Q., & Marshall, D. H. (1995). *Cochlear Implantation in the UK 1990–1994: Report by the MRC Institute of Hearing Research on the Evaluation of the National Cochlear Implant Programme*. Published by HMSO Books for the Department of Health, October 1995.

Torrance, G. W. (1986). Measurement of health state utilities for economic appraisal—a review. *Journal of Health Economics, 5*, 1–30.

Torrance, G. W., & Feeny, D. (1989). Utilities and quality-adjusted life years. *International Journal of Technology Assessment in Health Care, 5*, 559–575.

Walden, B. E., Demorest, M. E., & Hepler, E. L. (1984). Self-report approach to assessing benefit derived from amplification. *Journal of Speech and Hearing Research, 27*, 49–56.

Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-item Short Form Health Survey (SF-36), I: Conceptual framework and item selection. *Medical Care, 30*, 473–483.

Appendix: International Outcome Inventory for Hearing Aids (IOI-HA)

1. Think about how much you used your present hearing aid(s) over the past two weeks. On an average day, how many hours did you use the hearing aid(s)?

none	less than 1 hour a day	1 to 4 hours a day	4 to 8 hours a day	more than 8 hours a day
☐	☐	☐	☐	☐

2. Think about the situation where you most wanted to hear better, before you got your present hearing aid(s). Over the past two weeks, how much has the hearing aid helped in that situation?

helped not at all	helped slightly	helped moderately	helped quite a lot	helped very much
☐	☐	☐	☐	☐

3. Think again about the situation where you most wanted to hear better. When you use your present hearing aid(s), how much difficulty do you STILL have in that situation?

very much difficulty	quite a lot of difficulty	moderate difficulty	slight difficulty	no difficulty
☐	☐	☐	☐	☐

4. Considering everything, do you think your present hearing aid(s) is worth the trouble?

not at all worth it	slightly worth it	moderately worth it	quite a lot worth it	very much worth it
☐	☐	☐	☐	☐

5. Over the past two weeks, with your present hearing aid(s), how much have your hearing difficulties affected the things you can do?

affected very much	affected quite a lot	affected moderately	affected slightly	Affected not at all
☐	☐	☐	☐	☐

6. Over the past two weeks, with your present hearing aid(s), how much do you think other people were bothered by your hearing difficulties?

bothered very much	bothered quite a lot	bothered moderately	bothered slightly	bothered not at all
☐	☐	☐	☐	☐

7. Considering everything, how much has your present hearing aid(s) changed your enjoyment of life?

worse	no change	slightly better	quite a lot better	very much better
☐	☐	☐	☐	☐