Is Universal Newborn Hearing Screening More Efficient With Auditory Evoked Potentials Compared to Otoacoustic Emissions?

José Granell, Javier Gavilanes, Javier Herrero, Juan L. Sánchez-Jara, María J. Velasco, and Gonzalo Martín
Servicio de Otorrinolaringología, Complejo Asistencial de Ávila, Ávila, Spain

Introduction and objectives: Cost-effectiveness of universal newborn hearing screening programmes is under constant review. In this context, the aim of the present study is to evaluate the performance of brainstem response audiometry (BERA) compared to otoacoustic emissions (OAE) as screening tools.
Methods: Observational and retrospective study on a universal screening programme started in 1998. We perform a comparative analysis between 2 groups of newborns evaluated in consecutive periods of time. We analyze outcome measures of the programme as a measure of effectiveness, and dedicated resources to weigh the costs.
Results: We compare a group of 862 newborns from year 2003, screened with transient evoked OAE with a clinical device, with a group of 2300 newborns from years 2005 and 2006, screened with automated BERA. We find a statistically significant difference in the percentage of pass in the first step, favouring BERA (99.7% vs 91.8%; *P*<.0005). The median examination time with BERA was 276 seconds. Costs evaluation points to a progressively decreasing difference between both tools.
Conclusions: There are data indicating that BERA could be more cost-effective as initial screening tool. This advantage should be added to the already known more comprehensive evaluation of the auditory pathway, which could lead to the recommendation of its preferential use in auditory screening programmes.

Key words: Neonatal screening. Hearing loss. Otoacoustic emissions. Evoked auditory potentials.

Cribado universal de la hipoacusia neonatal: ¿es más eficiente con potenciales evocados auditivos que con emisiones otoacústicas?

Introducción y objetivos: La relación costeefectividad de los programmeas de cribado universal de la hipoacusia infantil es objeto de una constante revisión. En este contexto, el objetivo del presente trabajo es valorar el rendimiento de los potenciales evocados auditivos de tronco cerebral (PEATC) frente a las emisiones otoacústicas (EOA) como prueba de cribado.
Métodos: Estudio observacional retrospectivo sobre un programmea de cribado universal iniciado en 1998. Realizamos un análisis comparativo entre dos grupos de recién nacidos evaluados en períodos consecutivos, separados por una transición tecnológica. Analizamos los indicadores de resultados del programmea como medida de efectividad y los recursos invertidos para valoracion de costes.
Resultados: Se compara un grupo de 862 neonatos del año 2003, cribados con EOA transitorias con un equipo clínico, con otro de 2.300 neonatos de 2005 y 2006, cribados con PEATC automatizados. Encontramos una diferencia estadísticamente significativa en el porcentaje de paso en primera fase favorable a los PEATC (el 99,7 frente al 91,8%; p < 0,0005). El tiempo de exploración con PEATC tuvo una mediana de 276 s. La evaluación de costes apunta a una diferencia cada vez más ajustada entre ambas pruebas.
Conclusiones: Hay datos que indican que los PEATC podrían ser más coste-efectivos como prueba inicial de cribado. Esta ventaja se añadiría a la ya existente de evaluación más completa de la vía auditiva, y podría llevar a la recomendación de su uso preferente en los programmeas de cribado auditivo.

Palabras clave: Cribado neonatal. Hipoacusia. Emisiones otoacústicas. Potenciales evocados auditivos.

Correspondence: Dr. J. Granell.
Amado Nervo, 4, 8.º A. 28007 Madrid. España.
E-mail: jgranelln@seorl.net

Received March 7, 2007.
Accepted for publication January 28, 2008.

INTRODUCTION

The programmes for early detection of hearing loss in childhood by universal newborn screening have been developing over the past 15 years. The recommendations from 1993 by the expert group from the United States[1] remain essentially in place and most of the current programmes are based on them. It was assumed that auditory deficit in early childhood made language acquisition more difficult and altered the development of the children, their communication and learning skills, and in the long term, their social integration. It was also noted that it meant high costs, direct and indirect, by welfare, social and family resource consumption and that the degree to which this occurred depended on the severity of hearing loss and age of diagnosis. The incidence of neonatal hearing loss is estimated to be between 1/1000 and 5/1000 newly born.[2] The average age of diagnosis of hearing loss without perinatal screening programmes is more than 1 year.[3] The early detection and intervention, at 3 or 6 months old, improves language and speech development in comparison with interventions carried out after the first year of life.[4] The first screening programmes for hearing loss in newborns were developed in the sixties,[5] and were aimed at people with risk factors for hearing loss, present in 6%-8% of living newborns; in this population group the incidence of hearing loss rises up to 2.5%-5%.[6] However, it was soon recognized that the screening targeted at risk groups excluded 50% of newborns with hearing loss. Therefore, together with the availability of new and simple screening techniques, various agencies and groups of experts recommended universal screening,[7-9] for which there is a consensus. Nevertheless, at present there is no sufficient evidence to establish the long-term effectiveness of these programmes.

The methodology recommended initially in the various programmes was a sequential protocol based on the examination of newborns with otacoustic emissions (OAE) and/or brainstem evoked response audiometry (BERA). Undoubtedly, one of the most important consequences of the development and deployment for clinical use of the OAE, from a historical point of view, was precisely the introduction of universal screening programmes for infant hearing loss, as it was the first time that a technology became available which was valid and applicable to large groups of patients.[10] However, these protocols are undoubtedly influenced by technological developments. Indeed, the progressive introduction of these programmes originated a demand that led to the development of automated OAE equipment. Subsequently, the emergence of automated BERA equipment (aBERA) has led to the preferential use of the latter.[11] There is abundant literature comparing the 2 technologies regarding efficiency and effectiveness,[12] and which usually gives validity to both. Traditionally the advantages of the OAE are a greater speed of execution and a lower cost of equipment and consumables, although the work from which these conclusions are derived are not new.[13,14] The disadvantages traditionally attributed to the BERA are, by contrast, a higher completion time and greater cost. However, the aBERA have shown high percentages of "pass," even in very premature children,[15] and generally a lower rate of referral with programmes based on BERA is indicated, although no statistically significant difference has been demonstrated.[16,17]

The introduction of universal screening programmes for newborn hearing loss has been uneven.[18] The reasons are varied and their detailed analysis exceeds the objectives of this work. However, it is clear that one of the most important is the need to invest a significant amount of resources. In fact, the very story of the implementation of auditory screening programmes is closely linked to a continuing economic evaluation. Thus, it seems an appropriate time to review on the one hand, the necessity of screening for hearing loss, both from a medical and economic point of view, and the other hand, to seek more efficient study protocols. It is with this second goal in mind that we carried out the present work.

In our centre, testing for early detection of hearing loss in newborns with universal intent has been carried out since 1998. In the 10 years since then it has accumulated an experience that we might consider relevant in the practical operation of the screening programme. Since 2004 there has been a major change in both the technological and organizational aspects. The first is due to the creation of a centralized structure of the programme in the region, the second is a direct consequence of the first, the replacement of the OAE with aBERA as the technology used in the first phase of the screening. This trend provides us with a perspective from which we aim to analyze the technological aspects of the programme. The primary objective of this study is to compare the performance of aBERA with that of OAE in the universal screening programme for newborn hearing loss.

METHODS

We performed a retrospective observational study to compare the aBERA and the OAE on the universal screening programme for newborn hearing loss. The intention was to evaluate the efficiency of each test by analyzing key performance indicators of the programme, as a measure of their effectiveness, and thus make an estimate of costs.

According to the protocol, an initial screening test is carried out (stage 1) which is repeated in case of "fail" or "no pass" before one month of age (stage 2). In case of double failure, children are referred for diagnostic evaluation.

The 2 groups of this work correspond to 2 well differentiated periods. In what we shall called "stage 1" (from 1998 until April 2004), the programme depended on the resources of the otolaryngology service and the technology used in screening phases 1 and 2 was the OAE. An ILO 92® clinical device from Otodynamics (with ILO V.5 software) was used to explore whether or not OAE were evoked by transient stimuli. An average reproducibility over 85% in a test carried out under appropriate conditions (noise below 40 dB SPL and minimum stability of 90%) with 260 stimuli was required. The tests were conducted on a deferred basis after postpartum hospital discharge through subpoena

Key Indicators of the Auditory Screening Programme in the Study Groups[a]

	Group 1 (OAE)	*Group 2 (aBERA)*	*P*
Examinations assessed, n	825	2300	
Failure in stage 1	68 (8.2%)	8 (0.35%)	<.0005
Participation in stage 2	42/68 (61.76%)	4/8 (50%)	.705
Failure in stage 2	13 (30.95%)	0	.313
Deferral to diagnostic, %	1.5%	0	<.0005

[a]aBERA indicates automated brainstem evoked response audiometry; OAE, otacoustic emissions.

in the otolaryngology service. In the event of a failure, the test was repeated 1 month later with the same technology, and in case of a double failure, BERA was carried out with a clinical device (Centor C® from Racia Alvar). In newborns with risk criteria for hearing loss both OAE and BERA were carried out simultaneously in stage 1.

In 2004 the "Programme for early detection and comprehensive care of infant hearing loss" was started in our region.[19] In "stage 2" (from May 2005 until the present) aBERA are used in phases 1 and 2 (with an Accuscreen® screening device from Madsen, and software AccuLink® v1.11) and the same clinical BERA for diagnosis. Disposable adhesive headphones (echo couplers) are used for the screening and for the diagnostic, conventional earphones. Stage 1 is carried out before discharge from hospital after childbirth, in the maternity ward, where there is nursing staff dedicated to the programme.

The 2 study groups were extracted from each of the 2 stages identified. The patient selection was performed by temporary criteria and they were arbitrary. In "stage 1," the year 2003 was selected for the study, this was the last year in which the previous programme was in effect. A review and a registry in a database of examinations of that year was carried out. There is currently a computer database, which is centralized in the region and with restricted access, which is updated through the introduction of online data from the different centres that carry out the screening. From "stage 2" we used the data collected between January 2005 and December 2006.

The evaluation of resources consumed by explored case has been carried out generically. The reason is that the examination protocol is very different in the 2 stages, and a reliable comparison of costs is not feasible. Nevertheless, we shall note the factors that we believe have more influence in resource consumption. In addition, we have carried out an assessment of costs by analogy using a spreadsheet provided by the National Centre for Hearing Assessment and Management at the University of Utah.[20]

Statistical procedures for analytical comparison of proportions have been carried out; the χ^2 test was used in the event of meeting criteria of normality and the Fisher exact test in case non parametric tests were required. A *P* value less than .05 was considered significant. The SPSS 12.0 statistical package has been used.

RESULTS

Group 1, corresponding to the "stage 1," screening with OAE, consists of the 862 children born in our centre in 2003 on who the newborn auditory screening tests were carried out. This corresponds to 83.9% of newly registered live births (such was the coverage of the programme in that year). Group 2 consists of 2300 newborns explored between January 2005 and December 2006; the overall coverage of the programme in "stage 2" was of 96.4%. The average stay (standard deviation) after childbirth in our centre (data extracted by DRG for the year 2006) was of 3.5 (1.8) days. The deferred examinations are currently exceptional (less than 1%). There are also those born in private centres or in bordering regions without screening programme for hearing loss and which account for about 1.8% of the examinations.

Of the 862 newborns who form group 1, 37 are not measurable by lack of information or confusing information on the results of the screening test (although there is evidence of its implementation). The median age of completion of stage 1 is 11 days, with a sharply skewed distribution, with a minimum of 2 days and a maximum of 153 days. The 2300 newborns in group 2 are measurable. The average age of completion of stage 1 in this group is 1.8 days. The most significant quantitative results are presented in Table.

A total of 757 newborns in group 1 (91.75% of the 825 who could be evaluated) were given the classification of "pass" in stage 1. Participation in stage 2 was of 61.76%, since 26 of the 68 patients cited were "lost" (did not attend stage 2 having "failed" in stage 1). In stage 2 there was a 30.95% of failures, as 13 patients did "not pass." The latter were derived for diagnostic evaluation (this represents 1.5% of the 862 included in the programme). No child was confirmed as suffering hearing loss.

With respect to group 2, only 8 of 2300 babies did "not pass" in stage 1, which represents a rate of passes of 99.6%. This failure rate is still lower if we consider the year 2006 independently, where only 2 of 1202 children had to be re-examined (approximately 1.6/1000). The bilateral failure (both ears) in stage 1 occurred in 3 cases. Participation in stage 2 was of 4/8, that is 50%, at the expense of 4 patients who did not go to stage 2 in 2005 (all of them with unilateral failure in stage 1). No child required diagnostic evaluation.

Regarding the comparison of proportions of failures in stage 1 with both technologies, as noted above they are 68/825 for the OAE and 8/2300 for aBERA. The difference is statistically significant (*P*<.0005).

There is no reliable information on the costs of equipment. Clearly, the examination by BERA poses a greater consumption of consumables, since comparing the buds of the headset probe of the OAE (0.557 euros/unit) and the eventual replacement of filters (3.66 euros/unit) for every

40 ears explored, approximately, the use of adhesive disposable headphones (5.71 euros per pair) and the corresponding electrodes (3 per patient at 0.176 euro/unit) is required. This represents a cost per scan (1 patient, 2 ears) of 1297 euros for the OAE and 6238 euros for the BERA. The evaluation of resources used in the concept of staff is more complex: it requires a precise evaluation of the qualifications of the personnel employed and of the time dedicated to the programme. Clearly, the bulk of the time of the explorer is devoted to the first screening test. The examination time for both ears with OAE took an average of 151 (30) s. The time recorded on the aBERA examination device shows a median for both ears of 276 s (interquartile range of 369 s). In other words, the centralization rates for the examination time would be just over 4.5 min for the BERA, and 2.5 min for the OAE, and to this we should add the preparation time of the child and the registration of the data.

Regarding the generic assessment of costs, these vary according to what is assumed as the depreciation time of equipment, usage of consumables, percentages of "failures" expected, payroll, and qualifications of the personnel involved, and so on. For the different values introduced for these variables, we obtain costs for both technologies ranging between 10 and 15 dollars per newborn explored. We must bear in mind that with the failure rate in stage 1 which currently exists in our programme with aBERA, for the same volume of target population, a volume of about 8% less examinations are carried out in the screening phase with respect to the previous programme with OAE, and that a significant number of diagnostic assessments is avoided.

DISCUSSION

The resources needed to conduct a screening programme should take into account the number of tests to perform, that is the numerical value of the target population and the coverage to be attained (for which a minimum value of 95% has been defined). Our health area covers the whole province, with a single reference public hospital with maternity ward; this favours the exhaustiveness of the data. The number of births is about 1100 per year, although it is increasing. Approximately 1.5% of births are multiple (exclusively twins). The lack of infant ICU determines that a significant number of births with risk do not take place, and therefore this avoids a population with a known high risk of hearing loss. While it could be argued that this limited number of births facilitates the smooth operation of the programme; presumably in each context of health care, the resources devoted to it are proportional to the volume of the target population.

The differences in the numeric data of results between the 2 study groups are obvious, and clearly favourable to the current protocol based on aBERA, but there are several aspects that deserve to be commented. Firstly, regarding the arbitrary choice of the groups. The only information we can give to rely on the homogeneity of the groups is the fact of working with the population of infants rather than samples of it, managing performance indicators relating to this population; hopefully there will be no major population changes over such a limited period.

On the other hand, it is clear that the change of certain organizational aspects of the programme between "stage 1" and "stage 2" has been linked to an improvement in its efficiency. The coordination in the health service involves not only the development of an overall plan, but also an important effort of organizational and staffing resources, both of material and personnel to carry it out. Often in their inception, existing programmes in various centres have been possible without additional resources through the effort of convinced professional individuals; a continued programme is unsustainable in this way. The availability of specific screening equipment (screeners) is remarkable, as with these the test is faster and does not require validation by the specialist because the device itself gives the category of "pass"; the desirability of using automatic equipment for screening programmes does not require discussion.[21] Concerning the examination time, in our experience it seems comparable between OAE and BERA, because the fundamental condition is the tolerance of the child, that is, choosing the right time for testing, and not the specific technology employed. This impression must be assessed with caution since only 1 of the devices compared is automatic, and the circumstances under which the test is performed differ greatly in both protocols, as explained previously. The section of staff has a great importance in previously published assessments.[22] Differences in time dedicate to the programme are marked by the duration of examinations, as the work of registration, coordination, etc, is in principle independent of the technology used. Therefore, when the time of examination of OAE and BERA is similar, so does are the costs of both tests. The extrapolation of the estimated cost to our health care environment must be carried out with great caution in any case.

The management system of the database is extremely important,[23] since its volume can be massive. Centralization ensures continuity and access to data. Also, from an organisational point of view, testing before discharge after childbirth, ensures an adequate coverage and prevents subsequent travel and appointments at specific services. The detailed analysis of all these aspects, without doubt, would confirm a significant improvement in the efficiency of our programme, but does not provide any external validity to the results, since it is conditioned by idiosyncratic aspects of our health environment. To prevent any bias in this regard we will focus the discussion on the analysis of the technological aspects and not on the data conditional by organizational aspects of the programme.

Among the parameters of the programme, the one we consider more interesting is the percentage of failure in the first phase, because it is the one with most implications in the optimization of the programme. A high percentage of failure involves the repetition of tests in a significant number of infants (more so the larger the population is) and thus

an additional resource consumption, in addition to the risk of saturation of diagnostic services. It is one of the parameters that has been changed most to the point that it was previously found in an interval of percentage (8.2% in group 1) and is now found in an interval of per thousand (0.35% in group 2). Unlike previous studies,[16,17] this difference is statistically significant. This is so despite conducting OAE examination on a deferral basis, that is, theoretically circumventing a percentage of failures by early examinations, and also with a clinical device, which in principle has better specificity. Another important aspect to be assessed in the false positives of the screening (subjects who do not pass the first stage but who in subsequent phases are confirmed as having "normal hearing") is the concern generated in the parents. It is difficult to explain to patients the "philosophy" of the screening tests and their significance. For parents, the "not pass" of the first test raises suspicions of the existence of "deafness." Another interesting consequence is to prevent stage 2 from being left unattended, an old problem.[24] In this sense, it is worth remembering that efficiency is pointless if the programme is ineffective, and that in this case the effectiveness requires meeting the minimum coverage, both in stage 1 and stage 2. In our case participation in stage 2 has been low and comparable in both periods. The solution would be to strengthen "recovery" systems, an added problem to the organizational difficulties that the programme already has.

One last interesting aspect that should be commented in relation with the technological change is related to not conducting examinations prior to 38 weeks of gestational age (with the corresponding correction in premature cases). For the reasons already expressed, in our opinion this is an old criterion which we do not follow: while screening continues to be carried out as close to discharge as possible, prematurity is not a reason to differ it. We have not found faults which could be attributable to this cause. On the other hand, since all patients are explored with BERA, it is not necessary to consider the risk factors for newborn hearing loss with regard to postpartum screening.

By way of conclusions, in a context of, as yet, limited scientific evidence we have considered reviewing and offering our experience in the universal screening of infant hearing loss, which is, at least, extensive in time. The study presents obvious strengths and weaknesses. This is a work with constraints on its external validity, but it provides information on the practical implementation of a screening programme, and therefore is assessable regarding its effectiveness.

The question articulated in the title is too ambitious to be answered with a work with the methodological limitations of the current one, but the real object is not to answer it, but to raise it. There is evidence that suggests that screening in the first phase with aBERA could be more efficient than with OAE. In particular, in the present work we have found a statistically significant difference in the percentage of failure in stage 1, in favour of the aBERA. Confirmation of this data could lead to a greater cost-effectiveness ratio of the BERA with respect to the OAE. This confirmation would be made more precise with controlled prospective studies, which compared screening devices with both technologies. The added benefit of the BERA of recognizing hearing neuropathies should also be considered when assessing which is the most appropriate methodology at present for universal screening of infant hearing loss, particularly in newly created programmes.

Acknowledgement

To Luis Herrero Augustín, in memoriam. To Maribel Castro, nurse at the otolaryngology service, who has borne the brunt of the programme for so many years.

REFERENCES

1. Early identification of hearing impairment in infants and young children. NIH Consensus Statement. 1993;11:1-24.
2. Manrique M, Morera C, Morao M. Detección precoz de la hipoacusia en recién nacidos de alto riesgo. Estudio multicéntrico. An Esp Pediatr. 1994;40 Supl 59:11-45.
3. Martin JAM, Bentzen O, Colley JRT, Hennebert D, Holm C, Iurato S, et al. Childhood deafness in the European Community. Scan Audiol. 1981;10: 165-74.
4. Robinshaw HM. Early intervention for hearing impairment: differences in the timing of communicative and linguistic development. Br J Audiol. 1995;29:315-34.
5. Downs MP, Sterritt GM. A guide to newborn and infant hearing screening programmes. Arch Otolaryngol. 1967;85:15-22.
6. Parving A. Congenital hearing disability epidemiology and identification: a comparison between two health authority districts. Int J Ped Otorhinolaryngol. 1993;27:29-46.
7. Joint Committee on Infant Hearing. Position Statement. ASHA. 1994;36: 38-41.
8. Statement. European Consensus Development Conference on Neonatal Hearing Screening. Milan; 1998.
9. American Academy of Pediatrics. Task Force on Newborn an Infant Hearing. Newborn and infant hearing loss: detection and intervention. Pediatrics. 1999;103:527-30.
10. Kemp DT, Ryan S. Otoacoustic emission tests in neonatal screening programmes. Acta Otolaryngol Suppl. 1991;482:73-84.
11. Dort JC, Tobolski C, Brown D. Screening strategies for neonatal hearing loss: which test is best? J Otolaryngol. 2000;29:206-10.
12. Meier S, Narabayashi O, Probst R, Schmuziger N. Comparison of currently available devices designed for newborn hearing screening using automated auditory brainstem and/or otoacoustic emission measurements. Int J Pediatr Otorhinolaryngol. 2004;68:927-34.
13. Watkin PM. Neonatal otoacoustic emission screening and the identification of deafness. Arch Dis Child. 1996;74:16-25.
14. Davis A, Bamford J, Wilson I, Ramkalawan T, Forshaw M, Wright S. A critical review on the role of neonatal hearing detection in the detection of congenital hearing impairment. Health Technol Assesment. 1997;1(10).
15. van Straaten HL, Tibosch CH, Dorrepaal C, Dekker FW, Kok JH. Efficacy of automated auditory brainstem response hearing screening in very preterm newborns. J Pediatr. 2001;138:674-8.
16. Doyle KJ, Burggrraaff B, Fujikawa S, Kim J. Newborn hearing screening by otoacoustic emissions and automated auditory brainstem response. Int J Pediatr Otorhinolaryngol. 1997;41:111-9.
17. Korres SG, Balatsouras DG, Lyra C, Kandiloros D, Ferekidis E. A comparision of automated auditory brainstem responses and transiently evoked otoacoustic emissions for universal newborn hearing screening. Med Sci Monit, 2006; 12:CR260-3.
18. Thompson DC, McPhillips H, Davis RL, Lieu TL, Homer CJ, Helfand M. Universal newborn hearing screening: summary of evidence. JAMA. 2001;286:2000-10.
19. Programmea de detección precoz y atención integral de la hipoacusia infantil en Castilla y León. Consejería de Sanidad. Junta de Castilla y Leon; 2004.
20. National Center for Hearing Assessment and Management. Utah State University [cited March 1, 2007]. Available from: http://www.infanthearing.org/resources/cost/index.html

21. Grandori F, Sergi P, Pastorino G, Uloziene I, Calo G, Ravazzani P, el al. Comparison of two methods of TEOAE recording in newborn hearing screening. Int Journal of Audiology. 2002;41:267-70.
22. Weirater Y, Korth N, White KD, Downs D, Woods-Kerhner N. Cost analysis ot TEOAE-based universal newborn hearing screening. J Commun Disord. 1997;30:447-93.
23. Harrison WA, Dunnell JJ, Mascher K, Fletcher K, Vohr BR, Gorga MP, et al. Identification of neonatal hearing impairment: experimental protocol and database management. Ear Hear. 2000;21:357-72.
24. Folsom RC, Widen JE, Vohr BR, Cone-Weson B, Gorga MP, Sininger YS, et al. Identification of neonatal hearing impairment: recruitment and follow-up. Ear Hear. 2000;21:462-70.