RECENT ADVANCES

Universal neonatal hearing screening moving from evidence to practice

C Kennedy, D McCann

Recent technological advances have made feasible universal newborn hearing screening and therefore early detection of permanent childhood hearing impairment. Over the past three years, new information has been published on whether early intervention is beneficial, the possibility of harm arising from newborn screening, and its cost. Dramatic progress has been made in the large scale implementation of universal screening in many parts of the western world.

Any discussion of congenital permanent childhood hearing impairment (PCHI) needs to define the term at the outset. Bilateral PCHI = 40 decibels relative to hearing threshold level (dB HTL) is the target condition for which there is the strongest case for screening on the grounds that treating this condition early in infancy may reduce the average deficit in verbal as compared with non-verbal abilities from 25 intelligence quotient (IQ) points to 6 IQ points1 and, unless otherwise stated, is the condition to which “PCHI” refers in this article.

The hypothesis that early detection of PCHI produces worthwhile benefit in terms of improved speech and language provides the rationale for the screening of infants for the condition that has been undertaken in the United Kingdom for nearly half a century. Recent technological advances have made universal newborn hearing screening (UNHS) feasible and have not only led to a radical revision of screening methods but also opened the door to scientific evaluation of the benefit of early intervention. The enthusiasm of some advocates of UNHS in North America for the very early detection of even the mildest of hearing impairments is in contrast with a more sober systematic review of the topic on behalf of the US Preventive Services Task Force,2 which provides an excellent recent summary of the current state of knowledge.

The Current Topic review of this subject in this journal three years ago3 provides a background review of the issues, including a brief description of screening methods. The paragraphs that follow will focus principally on the many additional reports in the field that have been published since then.

There is new evidence on all four of the questions about translation of evidence into policy that were raised by Hall and Davis in their commentary on that review.4 Specifically, there is new information on whether early intervention is beneficial, the question of the possibility of harm arising from newborn screening, and its cost (but we will not review here the substantial number of papers on the financial aspects). Finally, dramatic progress has been made in large scale implementation of UNHS as a matter of public policy particularly in the United Kingdom and also in other parts of Europe and the United States.

EPIDEMIOLOGY

PCHI affects about 133 per 100 000 population (95% confidence interval (CI) 122 to 145); 112 are congenital and the remainder comprise progressive and acquired losses.5 6 Fortnum et al7 reported that the prevalence of PCHI of this degree continues to rise until the age of 9 years and may be as high as 205 per 100 000. The higher figure, if correct, suggested that progressive losses had previously been greatly underestimated. An accompanying editorial pointed out the need for longitudinal data to address this.8

This high prevalence, however, depended on inclusion of an adjustment for supposed under-reporting. The adjustment, arrived at by comparing reported figures from audiology departments and education departments and using a capture-recapture analysis, assumed that these sources were independent. This assumption was justified by reference to the small subpopulation of children who had received cochlear implants, but it is not likely to be true of the wider population of children with PCHI, particularly with respect to their contact with education and health services.

Reports on the prevalence of PCHI in large populations of screened infants in the United States, Australia, and the United Kingdom appear to support the more conservative figure of about 133 per 100 000.9 10 11 The fact that there are both progressive and acquired losses nevertheless indicates the need for systems to detect PCHI that develops after the newborn period. For a more detailed discussion of the epidemiological issues and their implications, the reader is referred to Fortnum’s own recent comprehensive review.12

Abbreviations: AABR, automated auditory brainstem response; PCHI, permanent childhood hearing impairment; TEOAE, transient evoked otoacoustic emission; UNHS, universal newborn hearing screening
ACCURACY OF UNHS AS ASSESSED IN INFANCY

Eight key criteria for assessing any screening programme were listed in the previous annotation on neonatal screening. One of these is the existence of a screening test that accurately identifies children who may have an impairment. The most important indicators of a screen’s accuracy are its sensitivity and specificity. In this case these are measures of the screen’s ability to identify infants with PCHI and those without PCHI respectively.

These tests can be assessed with respect to either the population that received a screening test or the larger population that was the target for screening for respective calculation of either the test or the programme sensitivity. The implications of a screening programme are also affected by the prevalence of the target condition. This is captured in additional measurements such as the positive predictive value, which is the probability that a case that screens positive will prove to have the target condition, and the positive likelihood ratio, a comparison of the odds of being a case after screening positive with the odds of being a case before screening.

In the case of PCHI, making these measurements is not straightforward. Estimates of sensitivity require ascertainment of all true cases, which is not possible until school age. They will include some cases of “progressive” loss and will not include audiology assessment of all those who screen negative. However, attempts to overcome these problems have been made (see next section). Many studies have avoided these problems by using a marker of process as a proxy measure of sensitivity. Incremental yield, the increase in the yield of the number of referrals of true cases, is an example of a proxy measure of sensitivity.

Estimates of specificity are easier to make but include the assumption that detection of “false positive” cases of mild, unilateral, or non-permanent hearing impairment is no more desirable than a false positive screen in a child with normal hearing.

Most programmes use a two step screen in which an infant who shows “no clear response” on the initial test undergoes a second test soon after the first, and, if the second test also shows no clear response, is referred for audiology evaluation (fig 1A–C). The principal neonatal screening tests are detection of transient evoked otoacoustic emissions (TEOAEs) and automated auditory brainstem response (AABR) testing. The screen consists of either repetition of one of these tests or sequential use of the two different methods either in all babies or, in other applications, only if the first test gives no clear response. A brief explanation of these methods is given in the previous annotation, and a more detailed description of their relative merits by Watkin, and a comparison of one and two stage screens formed part of the New York State UNHS demonstration project. Watkin’s recent review of these methods is a very clear account.

So how accurate is UNHS? Only two studies of the accuracy of UNHS were rated as being of good quality in a comprehensive systematic review of 864 articles on UNHS in 2000. The only controlled trial of UNHS reported that the yield of true cases of PCHI referred after 6 months of age was 32 per 100 000 in periods without and 94 per 100 000 in periods with UNHS, an increase of 62 cases per 100 000 target population. The other study of “good quality” according to the systematic review, was the New York State UNHS statewide “demonstration project”. This study reported a yield of 68 per 100 000 target population from UNHS. There was no comparison group to estimate the yield without UNHS. In that project, the mean (median) age of diagnosis was 3.5 (2.5) months for mild or moderate PCHI and 6.3 (3.8) months for severe PCHI. Developmental delay, non-attendance for follow up, and transient conductive losses delayed diagnosis until the second year of life in about 40% of cases.

The use of a reliable behavioural test of hearing, such as pure tone audiometry, in the entire population of children who had received UNHS would be necessary to estimate directly the sensitivity of UNHS to PCHI in infancy but the low prevalence of PCHI means that tens of thousands of audiograms would be required to produce reliable figures. No such study has been done except in a population of graduates of neonatal intensive care units. The higher risk of PCHI in this group reduced the number of audiograms required to produce reliable estimates of its prevalence.

These babies were neonatally screened using TEOAEs followed by AABR in a study in Seattle and later evaluated with visual reinforcement audiometry at the age of 8–12 months. “Hit rates” for the neonatal tests, in general, exceeded 80% when the false alarm rate was fixed at 20% and hearing impairment was defined as behavioural thresholds > 30 dB HL. This is a lesser degree of PCHI than has been used in the definition of the target condition for UNHS in the United Kingdom. As the hit rate increased in the Seattle study as the severity of hearing loss increased, these figures are encouraging evidence of the sensitivity of the neonatal screen in high risk infants.

The UK national protocol (www.nhsp.info) recommends TEOAE detection followed by AABR testing as a two step protocol.
screen. It recommends that all babies nursed on neonatal intensive care units for more than 48 hours should have both tests, whereas “well babies” should only proceed to AABR if TEOAEs were not detectable in one (or both) ears in the first step of the screen (fig 2A,B).

An alternative strategy would be to tailor the threshold for screening positive according to the age at screening and to the risk status of the baby. For example, one study classified detection of TEOAEs in one ear as a negative screen in babies screened less than 48 hours after birth but required bilateral TEOAE detection to fulfill criteria for screening negative in babies screened at a greater postnatal age.11 In the same study, amending the protocol to redefine a positive screen, in the babies at low risk13 only, as bilateral (rather than unilateral) failure on AABR testing (the second step of the screen) significantly reduced the overall false alarm rate from 1.15% to 0.56% without reducing the yield of true cases of PCHI from the screen.

That study provided a simple summary of the “accuracy” of this newborn screen as follows: of those screening positive on this protocol, a quarter (0.17% of those screened) had the target condition of PCHI, a quarter (0.17% of those screened) had non-targeted (non-permanent, mild, or unilateral) hearing impairments, and a half (0.34% of those screened) had normal hearing.22 The subsequent estimate of sensitivity from the same study suggests that, of those screening negative, less than 1 in 10 000 will subsequently prove to have a PCHI (see next section). These figures are very acceptable when screening for a condition of relatively high prevalence and exemplify achievable standards for accuracy of UNHS with which other implementations could be compared. They depend on setting the threshold for screening positive at a level high enough to keep specificity high but low enough not to compromise sensitivity.

UNHS programmes have also been introduced in other parts of Europe,23 the United States,24 25 and Australia.26 The effectiveness of many programmes in the United States has been limited by poor rates of screening and attendance for rescreen and for audiolog detection after “no clear response” on UNHS. A study in Rhode Island, the first state in the United States to establish a state wide programme of UNHS, has shown that this is heavily dependent on socio-economic and demographic variables: those with traditional Medicaid insurance were less likely to be screened or to be rescreened.26

**ACCURACY OF UNHS AS ASSESSED LATER IN CHILDHOOD**

Not all school age PCHI is detectable in infancy. As well as those cases in which negative screening in the neonatal period or later in infancy is misleading because deafness was, in fact, present, there are some cases in which the neonatal screen was correct but genetically determined or prenatally acquired PCHI subsequently progressively worsens with age (see section on epidemiology above), and others in which deafness is postnatally acquired. Consequently, the sensitivity of UNHS to childhood PCHI is lower than to infantile PCHI.

Follow up data derived from population based birth cohorts that were the target of UNHS are very sparse. Longitudinal study of the three year birth cohort enrolled in the Wessex trial12 now allows estimation of the sensitivity of the UNHS programme for all cases of PCHI still present in the population at the age of 7–9 years.23 Information from a number of professional and other sources indicated that the prevalence of PCHI in these children at 7–9 years of age was about 125 per 100 000 both in populations that were and were not the target of UNHS. In the half of the trial birth cohort that was the target for UNHS, sensitivity of UNHS to PCHI at 7–9 years was 56%. Of all such cases, 58% had been referred (from UNHS or other source of referral) before the age of 6 months. These figures compare well with the equivalent respective figures of 14% and 16% in the half of the trial birth cohort that was not the target for UNHS. Test sensitivity of UNHS was, nevertheless, high at 30/32 (94%) cases. In other words, only two of the 32 cases (eight per 100 000 target population) of PCHI in the population at 7–9 years of age had actually received UNHS and shown a clear response—that is, screened negative—yet subsequently been found to have PCHI. Watkin16 has reported 81% programme sensitivity and 94% test sensitivity of UNHS on the basis of longitudinal follow up to school age.

**DELETERIOUS EFFECTS OF UNHS**

To identify a single confirmed case of PCHI by UNHS, nearly 1000 infants will be tested, and between five and 40 will show “no clear response” on UNHS. It is therefore important to show a very low incidence of harmful effects from screening. Recent publications7 27 28 have confirmed the earlier reports29 30 showing positive effects and no increase in anxiety in the vast majority of families. The anxiety and other adverse effects seen in a very few can be avoided or minimised both by increasing the quantity and quality of presentation of information—for example, use of “no clear response” in preference to “fail”—and also by decreasing false alarm rates. False alarm rates can be reduced by adapting screening tools to the age at screening and risk status of the baby (see section on accuracy above).

**EFFECT OF EARLY INTERVENTION ON SPEECH AND LANGUAGE**

The deleterious effect of bilateral PCHI on speech and language is well established.31 32 Language quotients of left infants enrolled in the Colorado Home Intervention Program at a mean age of 26 months were reported to be substantially higher in those identified between birth and 6 months than in those identified subsequently (mean (SD) 79.0 (20.9) v 63.8 (19.3), p < 0.001), with little effect of the severity of PCHI on the magnitude of the difference.3 Those identified between 6 and 12 months fared no better than those diagnosed later. More recent reports from the same group have compared language quotients between children born in hospitals with and without UNHS programmes.33 They have summarised the results of their programme of research from 1994 to 2001.34 Mean language scores were 18–21 points higher and within the normal range for the neonatally screened group.

All or nothing nature of the reported benefit of early intervention reported from the Colorado programme has catalysed the introduction of UNHS in the United States, but the studies of these children are open to a number of criticisms and were rated as “poor quality” in the systematic review.7 In the earlier study, the later identified group was characterised by mothers of lower educational achievement and babies with more severe hearing loss and greater likelihood of non-verbal cognitive impairment and use of sign language, but the analysis did not allow simultaneous adjustment for all of these factors.

There was no information about the whole population of children with hearing impairment nor about children whose families had taken them out of the service before the study. In the more recent study, eligibility for the screened group was conditional on an assessment of language being available, the two groups were drawn from different hospitals at different times, selection of subjects and assessment of outcome were unblinded, and the number of and reasons for exclusion were not reported. Outcome measures relied heavily on parent reports, which are open to bias. Finally the fact that outcomes were assessed in the preschool years...
means that measurement of non-verbal IQ was difficult and likely to lack stability over time, and that it was too soon to assess catch up progress in verbal abilities among those treated later.

Benefits of early intervention were also reported in an uncontrolled sample of 112 children aged 5 years enrolled in the Diagnostic Early Intervention Program in Nebraska. Those with cognitive abilities less than 70 and non-participants in the programme were excluded. Age of enrolment in the programme accounted for 11.5% of the variance in vocabulary, and family involvement in treatment (assessed on a simple rating scale) accounted for 57% of the variance in a multiple regression analysis that also included non-verbal IQ. Rates of attrition from the study were not reported. Children with poor family involvement and late enrolment had poor speech and language outcomes, whereas good family involvement appeared to protect children from the deleterious effects of being enrolled late.

A population based study of all children in the state of Victoria, Australia with congenital PCHI reported on outcome at 7–8 years, an average interval of 6.1 years from diagnosis. All such children who were fitted with hearing aids by 4.5 years and without intellectual disability or a serious medical condition were eligible for inclusion. Outcomes were reported on 89 respondents from 132 eligible children from the entire three year birth cohort from a state population of

Figure 2 (A) Decision chart for UK national screening protocol: well babies. (B) Decision chart for screening protocol: babies in neonatal intensive care unit (NICU) for more than 48 hours.

The aim is to complete the screening process prior to discharge from hospital. If this is not possible an outpatient or community appointment to be arranged to complete the screening process. The target is to complete the screen by 4 weeks of age.
4.8 million. Respondents were similar to non-respondents, so the findings should be generalisable. The mean age at which hearing aids were fitted was 1.9 years. These children and their parents completed a battery of assessments of language, articulation, cognition, reading, adaptive functioning, health related quality of life, behaviour, and school functioning. The sample scored far (1.3–1.7 standard deviations) below the normal population on language scores. Outcomes were significantly worse on every outcome measure except physical health. Mean language and vocabulary scores, but not other outcome measures, fell sharply with increasing severity of hearing impairment adjusted for age at diagnosis and non-verbal IQ. By contrast, there was no significant correlation between age at diagnosis or age at treatment and any of the outcome measures. Family functioning did not correlate with any outcome measure, and socioeconomic factors were only weakly related to outcome measures. In a general linear model, estimated mean language and reading scores did not vary significantly by age of diagnosis adjusted for severity of PCHI and non-verbal IQ. Eleven children in the sample were diagnosed before the age of 6 months and their outcome was no better than the further 17 diagnosed between 6 and 12 months or the 55 diagnosed later.

This careful population based study thus showed the opposite result to that reported in children of similar age reported by Moeller. This must cast doubt on the generalisability of the two service based American studies. The main weakness of the Australian study is the relatively small numbers of children diagnosed in the first 6 months of life: although the study found no evidence of benefit in these children, the confidence intervals around this observation are wide.

A UK study of design similar to that of the Australian study will report in 2005 on speech and language and other outcomes in a cohort of over 100 children with PCHI born between 1993 and 1996 and a comparison group of normally hearing children. The children were drawn from the Wessex trial birth cohort, half of whom received UNHS, the Waltham Forest and Hillingdon birth cohorts that were the target for UNHS, and their two respective neighbouring districts that were not. Our preliminary analysis of language outcomes in this population, after adjustment for severity of PCHI, non-verbal intelligence, and maternal educational level, confirms significantly improved language skills at 7–10 years in children whose PCHI was confirmed before the age of 10 months compared with those in children whose PCHI was confirmed at a greater age (www.nhs2004.polimi.it).

In summary, there is some preliminary evidence for improved language outcomes following “earlier” intervention, but as yet this is inconclusive with respect to the threshold age by which intervention should begin and interrelations between age at intervention and other important explanatory variables such as severity of PCHI, non-verbal ability, family participation, and the nature of the intervention. Population based investigations of children of primary school age may help to clarify the benefit of early intervention on outcomes for children with PCHI.

IMPLEMENTATION OF UNHS AS NATIONAL POLICY IN THE UNITED STATES AND THE UNITED KINGDOM

Between 1993 and 2000, the number of hospitals in the United States that screened more than 90% of newborns for hearing loss increased from 11 to about 1000. Thirty two states have “mandated” UNHS by law, although the funding arrangements to support this are often problematic. Age of intervention is falling in successive birth cohorts, and the fall is associated with UNHS programmes. The comprehensive critical review commissioned by the Department of Health (DoH), published in 1997 and summarised in this journal, set in train the process, which has led to a national change in practice in the United Kingdom. In 2001, the DoH commissioned a national programme of UNHS and linked its implementation to a systematic evaluation of the first 23 sites with a view to involving the whole of the United Kingdom in the programme by April 2006. Major innovations in services are being developed to produce family friendly hearing services in which health education and voluntary and social services provide an equitable and seamless service for children with PCHI and their families. In particular, the Department for Education and Skills (DfES), working with the Royal National Institute for Deaf People, has devised new strategies for working with children under the age of 2 years for which they had not previously had a role, and Early Excellence Centres have been developed to promote multidisciplinary input to all stages of management.

Hall and Davis summarised the frustrations experienced in the late 1990s by those involved in trying to bring these changes about. They drew attention to the fact that the heavily centralised decision making process in the United Kingdom lacked clarity as to how the findings of a critical review may influence policy and that the children’s subgroup of the national screening committee lacked the authority to turn policy into action. There were no funds available to allow the questions raised by the systematic review to be addressed in a timely fashion, no defined way in which the views of the consumer, as represented by voluntary organisations, could be incorporated into policy making, and no commitment from the government to respond to the advice of professional committees.

Between 2000 and 2003, however, frustration seems to have given way to a stream of creative and sometimes radical change supported by (at least some) government funding. Visitors to the newborn hearing screening programme web site (www.nhs2004.polimi.it) cannot fail to be impressed by the wealth of information available, menus, submenus, downloadable pdf files for parents and professionals, PowerPoint presentations, information sheets in many languages, newsletters, links to other sites such as the DfES early support pilot programme site (www.espp.org.uk), and so on. A process of dynamic change is already well advanced. Web page feedback on the current performance of the programme included the following statistics: for the pilot sites up to 31 May 2003, 121 253 infants had been screened, 2341 referred from the screen, and 121 cases of bilateral hearing loss confirmed, a yield of 100 per 100 000, a figure remarkably similar to that reported in the Wessex Trial. The coverage rate from all sites is reported to be 96% overall. This is a quite astonishing achievement. This figure included data from two pilot sites that reported coverage of 98% using community based screening, for which the pre-existing evidence base was very slender. It remains to be seen whether community based screening is generalisable in the same way that hospital based screening has proved to be.

Introduction of UNHS has immediate implications for information (assisted by the recent introduction of NHS number allocation at birth), equipment, and the development of information technology. In addition, there are major implications for the development and quality of paediatric hearing services, education, and social services. Quality assurance will be based around peer review of services achieved by continuous audit and a three yearly programme of visits by practising professionals in health, education, and social services and parent representatives.

It seems probable that in five years from now, at least two thirds of all cases of PCHI in childhood will be diagnosed before the age of 6 months and that the profile of service provision to support children with such impairments will have undergone radical change compared with that available.
five years ago. Achievement of this will depend on financial support to meet the demand for new kinds of service that will increase as early detection following UNHS becomes the norm throughout the United Kingdom. Although it will be some time before reliable estimates of cost effectiveness are available, it seems likely that this change will have reduced the burden to families and education of speech and language impairments and their secondary consequences. This is something to celebrate.

ACKNOWLEDGEMENTS
We are grateful to Professor A Davis for commenting on an early draft of this paper, to the MRC Institute of Hearing Research for figs 1A, 2A, and 2B, and to the Wellcome Trust and Defeating Deafness.

We are grateful to Professor A Davis for commenting on an early draft of this paper, to the MRC Institute of Hearing Research for figs 1A, 2A, and 2B, and to the Wellcome Trust and Defeating Deafness.

References