Newborn hearing screening

Helfand M, Thompson D C, Davis R, McPhillips H, Homer C J, Lieu T L

Authors' objectives
To identify strengths, weaknesses, and gaps in the evidence supporting universal newborn hearing screening (UNHS), and to compare the additional benefits and harms of UNHS compared with those of selective screening of high-risk newborns.

Searching
MEDLINE, CINAHL, and PsycINFO were searched from 1994 to August 2001 for studies published in English; the search terms were reported. The reference lists of review articles were screened for additional studies and experts were contacted. Studies published before 1994 were identified from two systematic reviews published in 1996 and 1997.

Study selection
Study designs of evaluations included in the review
Controlled trials or studies on the accuracy, yield or harms of screening, were eligible for inclusion. Studies on the effects of screening, or of early identification and treatment, were also eligible for inclusion. Uncontrolled case series and case reports were excluded for the evaluation of the effects of screening, early identification and treatment. Studies that reported any information about adverse effects of screening or early diagnosis were also included.

Specific interventions included in the review
Studies of otoacoustic emission (OAE) and auditory brainstem response (ABR) were eligible for inclusion. Studies in which screening was conducted by physical examination, or with tests other than ABR or OAE, were excluded. Some studies used two-stage screening (e.g. an OAE followed by an ABR, or a repeated ABR), while others used a single stage. A screen was defined as positive if, based on whatever tests were done by the time of discharge from hospital, a referral for repeat testing or audioligic consultation would be recommended. Tests performed in the hospital during the birth admission were defined as 'screening' tests.

Reference standard test against which the new test was compared
The authors did not specify any inclusion criteria relating to the reference standard. Any testing done after the birth admission was defined as being part of an effort to establish the final diagnosis. One study used visual reinforcement audiometry, performed at age 8 to 12 months, as the reference standard.

Participants included in the review
Studies of population-based or hospital-based programmes in newborns were eligible for inclusion.

Outcomes assessed in the review
Studies of the effects of screening, or early identification and treatment, had to report data on language outcomes. No inclusion criteria relating to outcomes were specified for the other sections of the review. The outcomes reported in the review were the numbers-needed-to-screen (NNS), sensitivity, specificity, positive predictive values (PPVs), and language and communication skills.

How were decisions on the relevance of primary studies made?
Two authors reviewed titles and abstracts from the searches. Studies that met the inclusion criteria were then selected for inclusion in evidence tables. The authors did not state how many reviewers were involved in this process.

Assessment of study quality
Studies were classified as 'good', 'fair', or 'poor' using pre-specified criteria developed by the U.S. Preventive Services Task Force (see Other Publications of Related Interest no.1). Two authors rated the studies for methodological quality. The 13-member Task Force discussed the review, and examined and rated the quality of four key studies of early evaluation.
Data extraction
Two authors extracted data on population, test performance and outcomes; it was not stated whether the data extraction was conducted independently.

For studies of screening, the number of patients with a final diagnosis of bilateral sensorineural hearing loss (SNHL) divided by the number of neonates screened and its inverse, the NNS to identify one infant with bilateral SNHL, were calculated. Where possible, the NNS in high-risk infants was calculated.

For studies of accuracy, the sensitivity and specificity were calculated: sensitivity as the number of infants with hearing loss who screened positive divided by the actual number of infants with hearing loss, and specificity as the number of infants with normal hearing who screened negative divided by the total number of infants with normal hearing. Positive predictive values (the number of infants with hearing loss who screened positive and later proved to have permanent bilateral SNHL divided by the number of infants who screened positive) were also calculated.

Methods of synthesis
How were the studies combined?
The authors constructed a mathematical model of the likely benefits and harms of UNHS versus selective screening of a hypothetical cohort of 10,000 newborns. This was used to estimate prevalence, sensitivity and specificity, compliance, treatment effect size, and other model parameters from the included studies.

How were differences between studies investigated?
Differences between the studies were discussed in the text.

Results of the review
Ten studies of the yield of universal screening programmes, one of the accuracy of OAE and ABR in high-risk children, and eight of the effect of screening or early treatment on speech and language outcomes, were included.

Yield of screening (10 studies).
Studies of the yield of UNHS found that 666 to 1,422 newborns needed to be screened to identify 1 case of hearing impairment. The NNS were highest (1,422 and 925) in the 2 studies judged to be of good quality. Poorer quality studies reported higher yields due to the inclusion of infants who had mild, unilateral, or unconfirmed hearing loss. The NNS to identify 1 extra case ranged from 483 to 2,794 in low-risk newborns and from 75 to 208 in high-risk newborns.

Sensitivity and specificity (4 studies that also reported on the yield of screening and one additional study).
In the studies that also reported on the yield of screening, the sensitivity ranged from 85% (specificity 98.5%) to 95% (specificity 90%). Specificity was only reported for 2 studies and did not appear to be reported in the tables.

In one good-quality study the PPV was 6.7%. In healthy babies the PPV was 2.2%, meaning that only 1 of every 45 infants referred for further testing was found to have moderate or profound bilateral SNHL. For high-risk babies the PPV was 20%. There were no systematic differences in the performance of transient evoked OAE or ABR when used as an initial screening test.

One diagnostic accuracy study that used an appropriate 'gold' standard (visual reinforcement audiometry performed at age 8 to 12 months) was included. The sensitivity of OAE ranged from 80% (specificity not reported) for moderate hearing loss to 98% (specificity 80%) for profound hearing loss. For ABR, the sensitivity and specificity were 84% and 90%, respectively.

Effectiveness of early detection (number of studies unclear).
No studies compared the yield of UNHS with the yield of a comparable concurrent selective screening programme. In a trial of UNHS compared with clinical screening at age 8 months, UNHS increased the proportion of infants with moderate-to-severe hearing loss diagnosed by age 10 months (57% versus 14%), but did not reduce the rate of diagnosis after age 18 months.

Effects of identification and treatment prior to age 6 months on language and communication (8 studies).

No prospective, controlled study directly examined whether newborn hearing screening results in improved speech, language, or educational development. None of the 10 screening studies included in the other sections of this review reported the outcomes of treatment for infants identified as having hearing impairment. Eight additional studies provided some information on the effects of identification and treatment on language and communication. These found that intervention before age 6 months was associated with improved language and communication skills by age 2 to 5 years. However, these studies suffered from methodological flaws.

In a mathematical model based on the review, the authors estimated that extending screening to low-risk infants would detect 1 additional case before age 10 months for every 1,441 low-risk infant screened, and result in the treatment of 1 additional case before 10 months for every 2,401 low-risk infants screened. With UNHS, 254 newborns would be referred for audiological evaluation because of false-positive second-stage screening test results versus 48 with selective screening.

Authors' conclusions

Modern screening tests for hearing impairment can improve the identification of newborns with permanent hearing loss. However, the efficacy of UNHS to improve long-term language outcomes remains uncertain.

CRD commentary

This review was very difficult to follow, partly due to the wide scope of its objective. The review objective was clearly stated, but the inclusion criteria were not defined in sufficient detail. It was not possible to determine from these criteria exactly which studies were eligible for inclusion. In addition, the 'Results' section of the review referred to studies which were not listed as included studies.

A reasonable literature search was carried out, but the review was limited to studies published in English. Thus, the results of the review may be subject to language and publication bias. The studies were assessed for validity, but only very limited results of this process were reported; it is therefore difficult to determine the validity of the individual study results. Some details of the review process, which included appropriate steps to avoid bias, were reported.

The review was mainly based on a narrative synthesis, which appears appropriate given the differences between the studies. However, the narrative lacked structure and clear tabulation of the results. The authors also constructed a mathematical model based on the results of the review. No details of the construction of this model were given, so it is not possible to comment on its validity.

The results of the review should be interpreted with extreme caution given the limitations highlighted.

Implications of the review for practice and research

Practice: The authors did not report any implications for practice.

Research: The authors stated that longitudinal studies of UNHS are required to address the hypothesis that early intervention in terms of UNHS is a predictor of language acquisition, and to link short-term improvements to better function in later life. They also stated that better evidence about the effectiveness of UNHS is needed, which could be obtained via population-based studies that begin with inception cohorts and carefully report outcomes in all possible patient, as well as rates of loss to follow-up. Speech, language and scholastic achievement of deaf and hard-of-hearing children should be followed over time. Such studies should evaluate whether the long-term language outcome of deaf children improves as the age of identification increases.
Funding
Agency for Healthcare Research and Quality, contract number 290-97-0018.

Bibliographic details

Original Paper URL

Other publications of related interest

Indexing Status
Subject indexing assigned by CRD

MeSH
Hearing Loss /diagnosis; Hearing Tests /methods; Infant, Newborn; Neonatal Screening

AccessionNumber
12002008076

Date bibliographic record published
31/01/2005

Date abstract record published
31/01/2005

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.