Newborn Hearing Screening: A Summary of the Evidence

Diane C. Thompson, MS; Heather McPhillips, MD, MPH; Robert Davis, MD, MPH; Tracy L. Lieu, MD, MPH; Charles J. Homer, MD, MPH; Mark Helfand, MD, MS

Epidemiology

Each year, approximately 5,000 infants are born in the United States with moderate-to-profound, bilateral permanent hearing loss (PHL). Estimates of the incidence of moderate, severe, and profound congenital PHL among newborns range from 1 in 900 to 1 in 2,500.[1-7] Congenital PHL is associated with delayed language, learning, and speech development.[8-12] This delay is measurable as early as age 3 years[13] and has consequences throughout life. On average, deaf students graduate from high school with language and academic achievement levels below those of fourth-grade students with normal hearing.[14,15]

Diagnosis and treatment are often delayed until ages 1 or 2 in children with congenital PHL, particularly among children at low risk for PHL.[16-20] Current theory views auditory stimuli during the first 6 months of life as critical to development of speech and language skills.[21-23] Advocates of universal newborn hearing screening (UNHS) believe that earlier application of available therapies, such as speech and language therapy, amplification, and family support, could reduce or eliminate the gap in language skills between deaf and hearing children.[24,25] Selective screening of high-risk newborns is an alternative to UNHS. The incidence of PHL varies with race, birthweight, and other risk factors. Among infants in a neonatal intensive care unit (NICU), the risk of moderate-to-severe PHL is 10 to 20 times higher than in the general population.[26] In addition to NICU admission, the Joint Committee on Infant Hearing high-risk guidelines specify 4 other risk factors (Table 1).[27] From 10% to 30% of newborns meet these criteria, which can identify 50% to 75% of all cases of moderate-to-profound bilateral hearing loss.[2]

In 1995, the U.S. Preventive Services Task Force (USPSTF) found insufficient evidence to recommend UNHS.[28] They argued that, among low-risk infants, the prevalence of hearing impairment was very low, and substantial numbers of infants would be misclassified. They found that evidence for the efficacy of early intervention in patients diagnosed by screening was incomplete, but endorsed selective screening of high-risk newborns based on the higher prevalence of disease in this group.

From the Division of Medical Informatics and Outcomes Research and Evidence-based Practice Center, Oregon Health & Science University, and Portland Veterans Affairs Medical Center (Helfand), Portland, Oregon; University of Washington (Davis, McPhillips, Thompson), Seattle, Washington; Department of Ambulatory Care and Prevention, Harvard Pilgrim Health Care and Harvard Medical School (Lieu), Boston, Massachusetts; and the National Initiative for Children's Healthcare Quality (Homer), Boston, Massachusetts.

Address correspondence to: Mark Helfand, MD, MPH, Division of Medical Informatics and Outcomes Research, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Portland, OR 97201-3098. E-mail: helfand@ohsu.edu.

Reprints are available from the AHRQ Web site (www.ahrq.gov/clinic/uspstfix.htm), through the National Guideline Clearinghouse (www.guideline.gov), or in print through the AHRQ Publications Clearinghouse (call 1-800-358-9295 or e-mail ahrqpubs@ahrq.gov).

The USPSTF recommendations based on this evidence review can be found in Newborn Hearing Screening: Recommendations and Rationale (which precedes this chapter), available on the AHRQ Web site and through the AHRQ Publications Clearinghouse.

Table 1. Risk factors for sensorineural hearing loss in newborns*

1. NICU admission for 2 or more days.
2. Usher's Syndrome, Waardenburg's Syndrome, or findings associated with other syndromes known to include hearing loss.
3. Family history of hereditary childhood sensorineural hearing loss.
4. Congenital infections such as toxoplasmosis, bacterial meningitis, syphilis, rubella, cytomegalovirus, and herpes.
5. Craniofacial anomalies, including morphologic abnormalities of the pinna and ear canal.

*Joint Committee on Infant Hearing (JCIH) criteria for identifying infants at high risk for hearing loss.[27]

Since 1995, many health care professionals and Federal health care agencies have advocated for UNHS, which is now mandated by law in 32 states.[16,27,29] Is widespread support for UNHS now justified? To update the USPSTF recommendations, we critically reviewed recent evidence to identify strengths, weaknesses, and gaps in the evidence supporting UNHS.

Methods

We focused our literature search on key questions underlying the clinical logic behind screening for hearing impairment in newborns (Figure 1). The logic assumes that screening tests are accurate; that screening reduces delays in diagnosis and treatment; that earlier treatment results in better language function within the preschool period; and that this improvement in early language function will improve educational, occupational, and social function later in life. Moreover, for UNHS to be preferred over selective screening, the potential benefits of early detection and treatment must be realized in the subgroup of newborns who have no risk factors and would not otherwise be screened.

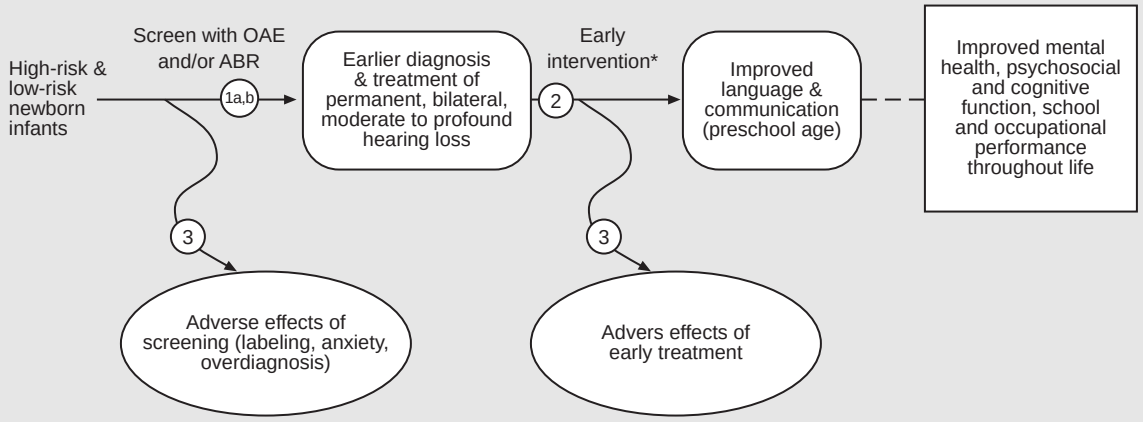
Search Strategy

We searched MEDLINE®, CINAHL®, and PsycINFO for relevant papers published in English from 1994 to September 2000, using the keywords *hearing disorders* and *infant or newborn* combined with terms for screening and relevant treatments, such as *early intervention, amplification*, and *American Sign Language* (see Appendix 1 for complete search strategy). The search was updated quarterly through August 2001. We also examined reference lists of review articles,[7,30-37] and queried experts. To identify articles published before 1994, we relied on systematic reviews published in 1996[28] and 1997.[18]

Study Selection, Data Abstraction, and Validity Assessment

The first author and at least 1 other author reviewed titles and abstracts of 864 articles and selected 340 articles as possibly relevant to 1 of the key questions. Of these, we selected to include in evidence tables (1) controlled trials; (2) reports on the accuracy, yield, or harms of screening using otoacoustic emissions (OAEs), auditory brainstem response (ABR), or both in the general newborn population; or (3) reports of the effects of screening, or of early identification and treatment, on language outcomes. Ten studies of the yield of universal screening programs,[24,38-46] 1 study of the accuracy of OAEs and ABR in high-risk infants,[47]and 8 studies of language outcomes[13,48-54] met these inclusion criteria. Two authors abstracted data on population, test performance, outcomes, and methodological quality from each included study. We classified each study as "good," "fair," or "poor" using prespecified criteria developed by the USPSTF for grading the internal validity of studies and the overall evidence for each link in the analytic

Figure 1. Newborn hearing screening analytic framework



*Hearing aids or other amplification, American Sign Language and/or English instruction, speech & language therapy, family education & support

Note: Numbers refer to key questions (KQ) as follows: (1a) Can UNHS accurately diagnose moderate-to-profound sensorineural hearing impairment? (1b) In UNHS programs, how many children are identified and treated before 6 months? (2) Does identification and treatment prior to 6 months improve language and communication? (3) What are the potential adverse effects of screening and early treatment?

ABR indicates auditory brainstem response; OAE, Otoacoustic emissions.

framework (Appendix 2).[55] When necessary, we sought additional information needed to apply the criteria from authors.

Synthesis

We constructed a mathematical model of the likely benefits and harms of UNHS versus selective screening of 10,000 newborns, estimating prevalence, sensitivity and specificity, compliance, treatment effect size, and other model parameters from the included studies. We used this evidence to prepare a technical report, summarized by the present article. The project team included investigators from the Oregon Health & Science University Evidence-based Practice Center, the University of Washington Departments of Pediatrics and Epidemiology, and 2 representatives from the USPSTF. The entire 13-member Task Force discussed the review, examined and rated the quality of 4 key studies of early intervention, and provided overall guidance.

Results

Can UNHS Accurately Diagnose Moderate-to-Profound Sensorineural Hearing Impairment?

Does UNHS Improve the Yield of Screening, Compared With Selective Screening of High-Risk Newborns?

Ten publications[24,38-46] provided information about the yield of UNHS and the performance of OAE and ABR in actual screening programs (Table 2). The 10 studies include 1 controlled trial, 4 state-based programs, and 5 hospital-based programs. Overall, screening detected 1 case of moderate-to-profound PHL for every 465 to 925 infants screened; from 779 to 2,794 low-risk, and 86 to 208 high-risk, newborns were screened to find 1 case. In these studies, the proportion of infants diagnosed to have significant hearing loss who had no risk factors varied from one-fourth to over two-thirds. Screening

Table 2. Studies of universal newborn hearing screening

Author, year	Description (quality rating)	Screening tests	# Screened/ # available (% screened)	Yield & NNS to find 1 case of bilateral PHL	# Positive screen (%) # follow-up %lost to follow-up	Definition of high-risk	# Low-risk identified/ # screened NNS	#High risk identified/ #screened NNS
Wessex, 1998[38]	Controlled, nonrandomized trial at 4 hospitals from 10/93 to 10/96 (Good)	TEOAE followed by ABR	21,279/25,609 (83)	23/21,279 925	342 (1.6) NR NR	NIH Criteria	7/19,555 2,794	20/1,724 86
Prieve, 2000[42]	State-wide demonstration project at 7 perinatal centers, 8 hospitals in New York (Good)	TEOAE followed by TEOAE or ABR in birth admission; TEOAE, ABR at 4-6 weeks (stage 2)	69,766/71,922 (97)	49/69,736 1422	4699 (6.5) 1st stage* 43.4	NICU infants	33/NR 2041†	52/NR 208†
Vohr, 1998[24]	Cohort from 8 maternity hospitals in Rhode Island from 1/93 to 12/96 (Fair)	TEOAE followed by: ABR (HR infants) TEOAE and ABR in 2-6 weeks (LR infants)	52,659/53,121 (99)	79/52,659 666†	5,397 (10.2) 1st stage 677 (1.3%) 2nd stage 4,575 15.2	NICU infants	61/47,529 779‡	50/5,130 103‡
Finitzo, 1998[41]	Cohort from 9 Texas hospitals from 1/94 to 6/97 (Fair)	ABR or TEOAE in birth admission followed by either ABR or TEOAE at 1-8 weeks	52,508/54,228 (97)	20/17,105 855† §	1,787 (3.4) 1,224 31.5	NR	NR	NR
Barsky-Firsker, 1997[43]	Hospital-based series at Saint Barnabas Medical Center, New Jersey 1/93 to 12/95 (Fair)	ABR by audiologists (One-stage)	15,749/16,229 (97)	NR	485 (3.1) NR NR	NICU infants	29/14,014 483	23/1,735 75
Watkin, 1996[44]	Hospital-based series at Whipps Cross Hospital, England (Fair)	TEOAE followed by TEOAE and ABR within 4 weeks	11,606/14,353 (81)	19/11,606 755	337 (2.9) 290 14	Risk factors and/or NICU infants	7/NR	13/NR
Mehl, 1998[40]	Cohort from 26 hospitals in Colorado from 1992 to 1996 (Poor)	19 ABR, 1 TEOAE, 6 ABR Follow-up screen not reported	41,796 /NR	NR	2,709 (6.5) 1,296 52.2	NR	NR	NR
Aidan, 1999[45]	Hospital-based series in Paris, France of infants in normal newborn nursery (Poor)	TEOAE in 48 hours; TEOAE within 4 weeks	1,421/1,727 (82)	2/1,421 711	238 (16.7) 123 48.3	hypoxemia, hyperbilirubin emia, FH	2/1,421 711	NR
Clemens, 2000[46]	Hospital-based series in North Carolina (Women's Hospital of Greensboro) from 7/98 to 6/99 (Poor)	ABR followed by re-test for fails (stage 1a) or ABR (stage 1b); outpatient ABR and diagnostic ABR (stage 2)	5,010/5,034 (99.5)	NR	103/5,054 85 17.5	NICU infants	NR	4/454 114
Mason, 1998[39]	Series of infants born at Kaiser, Honolulu from 3/92 to 2/97 (Poor)	ABR	10,372/10,773 (96)	12/10,372 864	415 (4.0) 362 12.8	NICU infants	5/8,971 1,794	7/1,401 200

* Reported different rates for misses and fails

† Includes mild, bilateral hearing loss

‡ Includes unilateral hearing loss

§ Data reported for 1996 only

Note: ABR indicates automated auditory brainstem response; FH, family history; HR, high-risk; LR, low-risk; NICU, neonatal intensive care unit; NIH, National Institutes of Health Consensus Development Conference; NNS, number needed to screen; NR, not reported; PHL, permanent hearing loss (moderate or worse); TEOAE, transient evoked otoacoustic emissions.

the low-risk or well-nursery population resulted in identification of 5 of 27 (18.5%) hearing-impaired infants in the Wessex trial,[38] 7 of 22 infants (32%) in the Whipps Cross study,[44] 8 of 15 (53%) hearing-impaired infants in the Hawaii study,[39] and 2 of 6 (33%) in North Carolina.[46] All of the U.S. studies that reported results for low-risk and high-risk groups separately defined "high-risk" as those who had NICU admission. The New York program examined differences between the NICU and well-baby nursery in detail. Overall, 1 in 884 newborns screened had bilateral hearing loss. In the NICU, where 90% of babies had other risk factors, 1 in 125 had hearing loss. In the well-baby nursery, where 30% had risk factors, 1 in 1,042 had hearing loss.[42, 56]

How Often Do False-Negative and False-Positive Screening Test Results Occur?

Either the OAE or the ABR is used as the initial test in screening. Criteria vary for defining a "pass" or "fail" on the initial screening test, and results are sensitive to equipment, the tester's training, and ongoing quality control. Most programs use a 2-stage approach, in which an infant who fails the initial test is retested and is referred for audiologic evaluation only if he or she fails the second test. False-negative and false-positive rates can be calculated based on the results of the initial test or on the overall results of the 2 stages of screening.

Three studies, a controlled trial of UNHS in Wessex,[38] a hospital based program in England,[44] and a report of statewide screening in Rhode Island,[24] provided some information about sensitivity and about the false-negative rate of the screening test (1-sensitivity). These studies reported the number of cases missed by screening and eventually diagnosed by other means, but they did not make a comprehensive effort to follow babies who had normal screening test results. The false-negative rates were 15%,[38] 6%,[24] and 11%.[44]

A nonrandomized, controlled trial of screening was conducted at 4 hospitals in the Wessex district of the United Kingdom.[38] Over 3 years, neonatal screening alternated with usual care every 4 to 6 months in 4 maternity hospitals. During the periods of neonatal screening, 21,279 of 25,609 eligible children (83%) had a transient evoked otoacoustic emission (TEOAE) test, followed by ABR testing for those with a positive TEOAE test. Newborns with positive ABR results were referred for audiological testing. All children in both the screened and unscreened groups received the existing screening program—the health visitor distraction test (HVDT)—at about 8 months of age. Children who did not pass the HVDT were also referred for audiological testing. The 2-stage screening protocol identified 23 of 27 (85%) infants who proved to have hearing loss upon follow-up.[57] In the Rhode Island program, 5 of 79 (6%) infants with bilateral hearing loss passed birth screening but were diagnosed between 5 and 22 months of age by other means.[24]

False-positive rates vary among centers and depend on the strategy and timing of testing. In the Wessex trial, the false-positive rate of the overall, 2-stage screening procedure was 1.5% (specificity=98.5%). Therefore, for every 1,000 normally hearing newborns who completed screening, 15 were referred for a full audiologic evaluation because of false-positive screening test results. The false-positive rate fell from 1.9% on the first postnatal day to 1.1% on days 2 through 4.

If an infant has a positive result on the screening test, how likely is it that the infant has hearing loss? Because the prevalence of congenital hearing loss is low, there are many more false positives than true positives; as a result, the positive predictive value (PPV) (number of infants with hearing loss and a positive test divided by the total number testing positive) is also low.

The programs in Table 2 used a 2-stage screening protocol, in which an infant who fails the initial test (an OAE or ABR) is retested, either in the hospital or as an outpatient within 12 weeks of discharge, and is referred for audiologic evaluation if he or she fails the second test. The PPV can be calculated for either the first stage or the second stage of screening. If both stages are performed while the infant is in the hospital, the PPV of the second-stage test determines who will be recalled for follow-up testing as an outpatient. In 1 good-quality study, the

overall PPV for the second-stage screening test was 6.7%.[38] In the well-baby nursery, the PPV was 2.2%, meaning that 1 of every 45 infants referred for outpatient audiologic evaluation eventually proved to have moderate-to-profound bilateral SNHL. For high-risk babies the PPV was 20% (18/90). None of the other studies in Table 2 provided sufficient data to determine the PPV for moderate-profound bilateral PHL.

The gold standard determination of permanent hearing impairment for validating results of screening tests is a combination of otolaryngological and audiological consultation, diagnostic ABR testing, and other electrophysiological testing.[58] These assessments have traditionally been performed after 6 months of age, but in some programs are done as early as 2 months of age. The reliability of the gold standard—behavioral and/or audiologic evaluation—increases with the age at which it is performed.

In the Wessex trial, the first audiometric examination was done when the babies were between 8 and 12 weeks of age. Of 158 infants who screened positive, 27 were diagnosed to have permanent sensorineural hearing loss; in 2 of these cases (7.4%), however, the "gold standard" diagnosis was wrong, and the babies proved to have normal hearing when re-examined at 4 months or 10 months of age.[57] This suggests that, even when formal diagnostic evaluation is performed following screening evaluations, 2 of 27 infants were misdiagnosed. Based on the 95% confidence interval (CI), the number of misdiagnosed cases could range from 0.24 to 6.5 infants. In another study,[44] 5 of 17 (29%) infants initially diagnosed to have moderate PHL were later found to have only mild hearing loss. None of the other studies in Table 2 followed patients long enough to determine when the audiometric diagnosis of PHL was incorrect.

In UNHS Programs, How Many Children are Identified and Treated Before 6 Months?

One indicator of the benefit of UNHS is the number of additional cases of significant hearing impairment that are diagnosed early. The Wessex trial did not directly compare the rate of early diagnosis and treatment for UNHS to that of selective screening of high-risk newborns. It did compare UNHS to no newborn screening, followed in both groups by HVDT at 8 months of age. In the Wessex trial,[38] for infants with moderate-to-profound hearing impairment, UNHS increased rates of referral to an audiologist by age 6 months (an increase of 51 per 100,000; CI, 7.4-94.0 per 100,000; *P*=.03), but did not increase rates of confirmation of diagnosis (*P*=.22) or initiation of management within 10 months (*P*=.08). Among those with moderate or severe hearing loss, however, screening led to highly significant increases in confirmation and management by 10 months of age. With UNHS, 13 of 23 (57%) children with moderate or severe impairment were diagnosed by 10 months, whereas during the period of time without UNHS, only 2 of 13 (14%) with hearing impairment were identified by then.[38] UNHS did not reduce the rate of diagnosis after 18 months, either overall (5/27 for UNHS vs 6/26 for the control group) or in the moderate-to-severe subgroup.

How much of the overall benefit in the Wessex trial can be attributed to screening low-risk infants? Compared with selective screening of high-risk newborns, universal screening diagnosed an additional 13 cases of moderate-to-profound bilateral hearing loss per 100,000 screened, or 1 additional case for every 7,692 screened, before 10 months of age.

Because the 9 other studies in Table 2 were uncontrolled, the effect on the timing of diagnosis, compared with selective screening of high-risk newborns, cannot be estimated. Some of them reported decreases in the age at diagnosis over time. During the 4 years of UNHS in Rhode Island, the mean age of hearing loss detection decreased from 13.3 months prior to implementing UNHS to 5.7 months by year 4.[24] In Hawaii, the average age of hearing-loss identification and fitting with hearing aids decreased as the percent of the population screened by UNHS increased. When 19% of the population was screened, the mean age of identification was 12 months and the mean age of amplification was 16 months. In the last year of the program, when 95% of the population was screened,

the average ages of identification and amplification were 3 months and 7 months, respectively.[59] In the Whipps Cross study,[44] performed in the United Kingdom, the mean age at amplification was 4.2 and 13.8 months for children with profound and moderate hearing loss, respectively.

Four of the 8 observational studies reported the mean age at the time of treatment. For hearing aid fitting, the mean age for all patients was 5.7 months,[24] 5.8 months,[39] and 7.5 months[60] in the 3 United States studies. In the Whipps Cross study,[44] performed in the United Kingdom, the mean age at amplification was 4.2 months for children who had profound hearing loss and 13.8 months for children who had moderate hearing loss. None of these estimates included children who, although screened, did not return for follow-up testing or treatment (that is, they were not calculated on the appropriate intention-to-treat basis). None of the studies reported information about the technical success of fitting, including how often the hearing aids were used.

The ages of diagnosis in the screening studies were all considerably earlier than those reported in a national survey.[20] The validity of this comparison is limited, because estimates from the screening studies did not include children who, although screened, did not return for follow-up testing or treatment (that is, they were not calculated on the appropriate intention-to-treat basis). These cases are included in surveys of the time lags in usual care, which are assessed retrospectively and therefore include children diagnosed at a later date, making it likely that the age at diagnosis will appear better in studies of screening than in surveys of usual care.

Does Screening Improve Language and Communication Skills?

No prospective, controlled study directly examined whether newborn hearing screening results in improved speech, language, or educational development. None of the state-based programs described in Table 2 reported the outcomes of treatment for infants identified to have hearing impairment.

One retrospective cohort study compared language performance in hearing-impaired children detected by UNHS (n=25) to unscreened children (n=25).[51] All study subjects were participants in the Colorado Home Intervention Program (CHIP), a program that provides hearing aids and home visits for children with hearing loss.[61,62] Children born after 1996 in a hospital that employed UNHS and who did not have significant cognitive delays were compared with children born since 1992 in hospitals without a UNHS program. Subjects were matched on degree of hearing loss (mild, moderate, moderately-severe, profound), cognitive quotient (CQ), and age at time of speech language evaluations. The 2 groups were similar in gender, ethnicity, presence of multiple disabilities, mode of communication, education of primary caregiver, and chronological age.

Mean scores for expressive, receptive, and total language were within normal range for the screened group and 18 to 21 points higher (P<.001) than the unscreened group (expressive language 82.9 [SE 3.7] vs 62.1 [SE 4.3]; receptive language 81.5 [SE 3.7] vs 66.8 [SE 4.0]; total language 82.2 [SE 3.3] vs 64.4 [SE 3.9]). A 20-point gap is more than 1 standard deviation lower than normal for age, which would indicate that a child with average intellect would have the language abilities of a child who had an IQ of 80. Children identified prior to 6 months (whether in the screened or unscreened group) had a smaller gap between language development and cognitive ability than children identified after 6 months. Language development was within normal range for 56% of the screened group compared to 24% of the unscreened group.

While this study used relevant, validated measures of language outcomes and controlled for several important potential confounders, the creation of the study groups and description of the patients limited the conclusions that could be drawn. Eligibility for the screened group was determined by the availability of an assessment of language outcomes at 2 to 4 years of age. Because the groups were drawn from different hospitals and time periods, factors other than exposure to UNHS might have influenced outcomes. Selection of subjects and assessment of outcome were unblinded, and neither

the number of excluded subjects, nor the reasons for exclusion, are reported.

Does Identification and Treatment Prior to 6 Months Improve Language and Communication?

Older studies comparing early-identified to late-identified children with impaired hearing consisted of clinical series or case-control studies of highly selected patients, with heterogeneous causes of hearing loss, inconsistent definitions of early diagnosis, incompletely defined treatment regimens, and inadequate control for potential confounders.[18,28] None of these older studies examined the outcome of delayed diagnosis in children who have no risk factors for hearing impairment at birth.

Table 3 summarizes methodologic aspects and results of 8 retrospective cohort studies from 3 intervention programs.[13,48-54] All of these studies used standardized receptive and expressive tests to evaluate speech and language skills in preschool children, and all reported statistically significant associations between the age at the time of diagnosis and language development at 2 to 5 years of age. Adjusted mean scores for expressive and receptive language were 15 to 20 points higher in groups of children identified and treated early compared to the later identified groups.

Five studies reported speech and language results for children enrolled in the Colorado Home Intervention Program.[13,48-50,52] The most widely cited of these studies compared 72 hearing-impaired children identified prior to 6 months of age to 78 hearing-impaired children identified after 6 months.[49] After adjustment for cognitive function, children whose hearing losses were identified by 6 months of age demonstrated significantly better receptive, expressive, and total language scores than children identified after 6 months of age. For children with normal cognitive abilities, this language advantage was found across all test ages, communication modes, degree of hearing loss, and socioeconomic strata. The children identified before 6 months of age had language scores at or near their cognitive test scores, whereas children identified after 6 months of age performed, on average, 20 points lower on language scores than cognitive scores. Children with low cognitive abilities (CQ <80) experienced a smaller improvement in total language, but no statistically significant improvement in receptive and expressive language abilities.

The groups of early and late diagnosed children differed: late-identified children were more likely to be cognitively impaired, to have severe or worse hearing loss, to use sign language, and their mothers were less likely to have finished high school. The statistical method used in the analysis did not simultaneously adjust for more than 2 factors and may not have removed the influence of these differences. Additionally, the study did not provide data on dropout rates in the 2 groups, and outcome assessments were not masked.

Another CHIP study evaluated factors related to expressive language development in a group of 113 deaf and hard-of-hearing children.[52] It reported that expressive vocabulary was higher with increased age, increased CQs, identification of hearing loss by the age of 6 months, and having a hearing loss as the only medical condition.

Additional evidence for the effect of early identification and treatment was provided by a retrospective study of 112 children enrolled for at least 6 months in a diagnostic early intervention program in Nebraska.[53] After adjustment for family involvement, degree of hearing loss, and nonverbal IQ, children enrolled prior to 11 months had stronger vocabulary and reasoning skills than children enrolled at later ages. At age 5, family involvement accounted for 57% of variance in vocabulary, and age of enrollment accounted for 11.5%. In 1 study, a retrospective series of 80 children in a home intervention program in Washington State,[54] early enrollment was associated with better language skills at 3 years of age. The relevance to newborn screening is low because only 9 subjects were enrolled before 12 months of age.

The studies in Table 3 had several important limitations. The study populations were composed of convenience samples. That is, the studies

Table 3. Cohort studies reporting language outcomes

Study, year (quality)	Selection of subjects	Comparability and maintenance of early vs late groups	Adjustment for cofounders	Results
Studies from Colorado Home Intervention Program (CHIP)				
Apuzzo, 1995[48] (FAIR)	Convenience sample of 69 high-risk infants diagnosed between 2 and 25 months of age. Children with severe cognitive delay were excluded.	Late-identified group was more likely to have severe to profound hearing loss (65% vs 50%). No report of attrition or follow-up rates	One-way ANOVA did not adjust for SES, family involvement, or other potential confounders.	At 40 months of age, infants identified before 2 months of age had higher mean Minnesota Child Development Inventory (MCDI) scores for expressive language (P<0.01).
Yoshinaga-Itano, 1998[50] (POOR)	Convenience sample of 40 high-risk infants, divided into those identified and treated before 6 months of age (n=15) and those treated after 18 months (n=25). Children with severe cognitive delay were excluded (DQ<60)	Late-identified group was more likely to have severe to profound hearing loss (52% vs 47%). No report of attrition or follow-up rates.	Gender, severity of hearing loss, cognitive function, and other disabilities were examined in 2-way ANCOVAs, not in a multiple regression (no simultaneous adjustment for multiple confounders).	At 40 months, infants identified before 6 months of age had better adjusted mean MCDI scores for expressive language (81.1 vs 64.3, P<0.05) and receptive language (84.4 vs 70.1, P<0.05).
Yoshinaga-Itano, 1998[13] (POOR)	Convenience sample of 82 infants, 19 to 36 months of age, with mild to profound PHL, divided into those identified before 6 months of age (n=34) and between 7 and 18 months of age (n=48). Early group identified by high-risk registry; late group by usual care. Children with severe cognitive delay were excluded (DQ<60).	Late-identified group was more likely to have severe to profound hearing loss (77% vs 42%). No report of attrition or follow-up rates.	Gender, severity of hearing loss, cognitive function, and other disabilities were examined in 2-way ANCOVAs, not in a multiple regression (no simultaneous adjustment for multiple confounders).	At 26 months, infants identified before 6 months of age had better adjusted mean MCDI scores for expressive language (76.2 vs 56.6, P=0.001), receptive language (82.1 vs 58.3, P=0.002), MacArthur CDI adjusted mean receptive vocabulary (200 vs 86.4, P<0.001), and expressive vocabulary (117 vs 54, P<0.03).
Yoshinaga-Itano, 1998[49] (POOR)	Convenience sample of 150 children 13 to 36 months of age with mild to profound PHL, divided into those identified before (n=72) or after (n=78) 6 months of age. The number of low-risk infants and the role of UNHS in identifying subjects are not described. Selection bias is likely because the design probably excluded infants who were diagnosed to have hearing loss but did not enter the program, or who entered, but were lost to follow-up.	At baseline, compared groups differed in some demographic characteristics and in the proportion of subjects with cognitive impairment and severe to profound hearing loss (CQ <80, 29% early group vs 56% late group; severe to profound hearing loss 34% early group vs 46% late group). No report of attrition or follow-up rates.	There was stratification by CQ (<80 vs >80). Other covariates (gender, minority status, maternal education level, Medicaid status, severity, mode of communication, other disabilities) were examined singly in 2-way ANCOVAs.	At 13 to 36 months, adjusted mean MCDI receptive language LQ was higher for those identified before 6 months (79.6 vs 64.6, P<0.001). Mean MCDI expressive LQ was higher (78.3 vs 63.1, P<0.001) and total language (79 vs 64, P<0.001) was higher in early-identified group. No differences in LQ among 4 age of identification levels in late-identified group.

continued

Note: ANCOVA indicates analysis of covariance; ANOVA, analysis of variance; CQ, cognitive quotient; DQ, developmental quotient; LQ, language quotient; PHL, permanent hearing loss; SE, standard error; SES, socioeconomic status; UNHS, universal newborn hearing screening

Table 3. Cohort studies reporting language outcomes (continued)

Study, year (quality)	Selection of subjects	Comparability and maintenance of early vs late groups	Adjustment for cofounders	Results
Studies from Colorado Home Intervention Program (CHIP)				
Yoshinaga-Itano, 2000[51] POOR)	Children born in a hospital with a UNHS program in effect at time of birth (n=25) were compared to children born in a hospital without a UNHS program (n=25). All subjects had been enrolled in CHIP program. Eligibility for the screened group was determined by the availability of an assessment of language outcomes.The creation of the study groups and description of the patients limited the conclusions that could be drawn.	The exposure was birth at a hospital with a UNHS program, not age of identification. Because the groups were drawn from different hospitals and time periods, factors other than exposure to UNHS might have influenced outcomes. Selection of subjects and assessment of outcome were unblinded, and neither the number of excluded subjects, nor the reasons for exclusion, are reported.	Pairs matched on age of testing (9-59 months), degree of hearing loss (mild, moderate, moderately severe, profound), and CQ.	Mean scores for expressive, receptive, and total language were within normal range for the screened group and 18 to 21 points higher ($P<.001$) than the unscreened group (expressive language 82.9 [SE 3.7] vs 62.1 [SE 4.3]; receptive language 81.5 [SE 3.7] vs 66.8 [SE 4.0]; total language 82.2 [SE 3.3] vs 64.4 [SE 3.9]). Language development was within normal range for 56% of the screened group compared to 24% of the unscreened group
Mayne, 2000[52] (POOR)	Convenience sample of 113 children 24 to 73 months of age, divided into those diagnosed before and after 6 months of age. The number of low-risk infants and the role of UNHS in identifying subjects are not described. Overlap of sample with previous CHIP studies was not reported.	Demographic comparisons of the groups were not reported. No report of attrition of follow-up rates.	Regression analysis adjusted for degree of hearing loss, mode of communication, other disabilities, parents' hearing, cognitive quotient, mother's education, ethnicity, SES	At 24 to 36 months, age at diagnosis explained 23% of the variance in expressive language scores.
Studies from other programs				
Moeller, 2000[53] (FAIR)	Convenience sample of 112 5-year-olds who completed the Diagnostic Early Intervention Program in Lincoln, Nebraska. Children with non-verbal IQ <70 and those who did not participate in program through age 5 were excluded. The number of low-risk infants and the role of UNHS in identifying subjects are not described. Outcome assessments were made pre- and post-intervention.	Not reported. No report of attrition or follow-up rates. Early identified children may have more opportunity to drop out, although differential drop out may be less of a problem at 5 years than in studies assessing closer to enrollment.	Multiple regression analysis adjusted for family involvement, degree of hearing loss and non-verbal IQ.	At age 5, family involvement accounted for 57% of variance in vocabulary and age of enrollment accounted for 11.5%. Adjusted mean vocabulary and reasoning scores were within normal range among children enrolled prior to 11 months but were lower for later-identified children.

continued

Note: ANCOVA indicates analysis of covariance; ANOVA, analysis of variance; CQ, cognitive quotient; DQ, developmental quotient; LQ, language quotient; PHL, permanent hearing loss; SE, standard error; SES, socioeconomic status; UNHS, universal newborn hearing screening

Table 3. Cohort studies reporting language outcomes (continued)

Study, year (quality)	Selection of subjects	Comparability and maintenance of early vs late groups	Adjustment for cofounders	Results
Studies from other programs				
Calderon, 2000[54] (FAIR)	Cohort of 80 children with profound hearing loss enrolled in Early Child Hearing Intervention (ECHI) in Seattle, Washington. Children with developmental delay were excluded. Cohort grouped by 3 levels by age of entry into program: <1 year (n=9), 12-24 months (n=39), >24 months (n=32). The method of sampling is not described, but the design excluded patients who entered the program but did not graduate.	Not reported. Late diagnosed group had less severe-to-profound loss (36% vs 66%). Overall loss to follow-up not reported. Because the early-diagnosed group were in the program longer, they had more opportunity to drop out, so a differential loss to follow-up is likely.	Controlled for degree of hearing loss, degree of outcome impairment that was present upon entry into program (baseline test levels).	At 3 years, age at entry to program explained 43.5% of the variance in receptive language and 49% of the variance in expressive language. Children treated before 2 years had better outcomes than those treated after 2 years. Only 3 children entered the program prior to 6 months of age.

compared children who were identified early and late by means other than UNHS, rather than children whose age at identification and enrollment was determined primarily by whether or not they were screened. None of the studies had clear criteria for inclusion, none had blinded assessments, and all selected children for inclusion based on the availability of a language assessment between ages 2 to 5. This could introduce bias: early-identified children who remained in the program may have had better results than early-identified children who were not available for follow-up. Because of these limitations, selection bias cannot be confidently ruled out as an explanation for the findings. Moeller[53] found family involvement an important contributor to language development. Since other studies did not adjust for this factor, they may have overestimated the association of early enrollment with language development. None of the studies provides information on attrition or follow-up rates. The USPSTF rated the strength of evidence linking early treatment with improved language function "inconclusive" and the quality of evidence as "fair/poor."

What are the Potential Adverse Effects of Screening and of Early Treatment?

Screening

Potential adverse effects of false-positive screening tests include misdiagnosis, parental misunderstanding and anxiety, and unfavorable labeling. As noted earlier, the "gold standard" determination of PHL is imperfect; in expert hands, as many as 7% of infants diagnosed to have PHL may eventually prove to have normal hearing. The frequency of misdiagnosis in everyday practice settings has not been studied.

Another potential adverse effect of screening is parental anxiety. In the Wessex trial, parents whose babies were screened had similar anxiety and attitudes to parents in the unscreened group. It should be noted that, before screening was done, parents in the screened group received information about the benefits of early identification and gave informed consent for the procedure. In the Whipps Cross hospital study,[63] among parents whose infants failed the initial screen and received a second test, 2 of 57 (3.5%) reported they were very worried. In a survey at a regional hospital in Logan, Utah (n=169), parents indicated acceptance of newborn screening for their infants: 98.2% of parents said they would give permission for screening, 95.3% would prefer screening even if the baby failed, and 84.9% felt that anxiety caused by failing a screening test would be outweighed by the benefits of early detection.[64] In another survey of non-NICU infants who failed hospital screening, 46 of 49 (94%) parents of infants who had false-positive screening test results approved of UNHS. However, 4 of 49 (8%) mothers said they treated their child differently (eg, spoke louder or clapped their hands), and 7 of 49 (14%) reported "lasting anxiety" after the second screening exam even though hearing was normal.[46] No study attempted to assess the effect of parental anxiety or changes in parental behavior on infants' development or on the parent-infant relationship.

Treatment

The harms of early intervention have not been adequately studied. As noted by the previous USPSTF, differing ethical and philosophical attitudes about deaf awareness and culture have led to controversy about the content of early interventions.[28] The argument for early intervention is based on the prevailing theory of language development, which holds that early auditory input is an important precursor of language development. An opposing viewpoint expressed in the literature is that, during infancy, nonverbal communication, joint attention, shared experiences, and mutual understanding are more important precursors of language development than are hearing speech and forming sounds. Proponents of this view theorize that early intervention could harm infants because it leads parents to focus on "means of communication the child has the least prerequisites for" and on the baby's disability instead of its competencies.[65] Because there are no randomized trials of different management strategies, it is impossible to assess the merits of these concerns.

Table 4. Benefits of screening a hypothetical cohort of 10,000 newborns for moderate-to-profound PHL

Benefit and relevant factors	Probability or effect size	UNHS	High-risk screening
Assumptions*			
Proportion high risk	0.2		
Prevalence			
High-risk group	0.008		
Low risk-group	0.0008		
Miss rate for UNHS (proportion not screened in hospital)			
in high risk	0.1		
in low risk	0.05		
follow-up rate for misses	0.9		
Miss rate for high-risk screening			
in high risk^	0.2		
followup rate for misses	0.75		
Sensitivity of 2-stage screening	0.85		
Specificity of 2-stage screening	0.97		
Compliance with follow-up	0.9		
Accuracy of diagnostic ABR			
Sensitivity^	1		
Specificity	0.995		
Proportion of low-risk diagnosed before 10 months without screening^	0.35		
Treated before 1 year	0.6		
Results			
Number of infants screened		9,400	1,600
Cases diagnosed before 10 months		17	12
Cases treated before 10 months		10	7
out of total		22	22
"False Positive" screening tests		254	48
Normal infants incorrectly diagnosed to have PHL at first post-hospital audiologic examination		1	0
NNS to diagnose 1 case		584	173
NNS to diagnose 1 additional case before 10 months		1,441	
NNS to treat 1 additional case before 10 months		2,401	

*Base case assumptions are derived from the studies in Table 2, except for those marked with ^.

Note: ABR indicates automated brainstem response; NNS, number needed to screen; PHL, permanent hearing loss; UNHS, universal newborn hearing screening

Summary of Benefits and Harms

Table 4 summarizes the benefits and harms of UNHS and selective screening in a hypothetical cohort of 10,000 newborns. With UNHS, an additional 7,800 screening tests would be done, resulting in the diagnosis of 6 additional cases of moderate-to-profound hearing loss diagnosed before 10 months of age. Of these, 3 additional cases would be treated before 10 months of age. Thus, the number needed to screen (NNS) to detect 1 additional case before 10 months would be 1,441, and the NNS to treat 1 additional case before 10 months would be 2,401. With UNHS, 254 newborns would be referred for audiological evaluation because of false-positive second-stage screening test results, versus 48 for selective screening. Of these, 1 would be falsely diagnosed to have PHL.

Of the 6 additional early-diagnosed, low-risk newborns, how many would actually benefit from early treatment? The data needed to estimate this—the probabilities of a poor language outcome with and without early treatment—are not known. To use a hypothetical example, if 50% of low-risk newborns would have poor language ability if diagnosed after 10 months, and early intervention

reduced this by 50%, then the NNS to prevent 1 additional case of delayed language acquisition would be 6,771.

Comment

Table 5 summarizes the evidence for each of the major assumptions underlying the case for UNHS. Several gaps in information about UNHS effectiveness remain. It is clear that modern screening tests for hearing impairment can improve identification of newborns with PHL, but as many as 10% of newborns with normal or temporarily impaired hearing will require a second screening test. From 1% to 3% of newborns will be referred for audiological assessment; over 90% of those referred are false positives. The consequences of these false alarms have not been adequately evaluated, nor has the reliability of audiological and behavioral assessment—the reference standard used to make the definitive diagnosis of hearing impairment—been adequately assessed in the setting of UNHS.

A clearer picture of the consequences of delayed diagnosis in low-risk newborns would strengthen the case for universal screening. Epidemiologic studies indicate that language development is often delayed in children with congenital hearing impairment and that the diagnosis of hearing impairment is often delayed. However, no study has examined language development in infants who were diagnosed at 1 or 2 years of age and who had no other disabilities and no risk factors for hearing impairment at birth.

Because the frequency and severity of poor language outcomes in this group is uncertain, only adequately controlled trials can establish the efficacy of early intervention. Several retrospective cohort studies show that, by 2 to 4 years of age, children who have had hearing aids and other therapy in the first 6 months of life had better language skills than

Table 5. Strenth of evidence for universal newborn hearing screening

Key question	Evidence code	Quality of evidence
1a. Can UNHS accurately diagnose moderate-to-profound sensorineural hearing impairment? *OAE and ABR are highly accurate screening tests for congenital PHL (sensitivity 84%, specificity 90%).*	II-1, II-2	Good: One controlled trial measured the predictive value of a positive test result (6.7%), and a good quality cohort study measured sensitivity and specificity against an independent gold standard.
1b. In UNHS programs, how many children are identified and treated before 6 months? *UNHS increases the chance that diagnosis and treatment will occur before 6 months of age. UNHS increases early identification between 19% and 42% over selective screening in high-risk children.*	II-1, II-2	Good: One controlled study in the United Kingdom and 1 cohort study in the United States reported the frequency of treatment before 10 and 6 months, respectively. Other studies did not provide sufficient information, and none included patients who, although screened, were diagnosed and treated late because of loss to follow-up. However, no controlled trials of UNHS versus selective screening have been done.
2. Does identification and treatment prior to 6 months improve language and communication in infants who would not be diagnosed that early in a selective, high-risk screening program? *Evidence is inconclusive.*	II-2, II-3	Fair/Poor: Studies have selection bias and baseline differences between compared groups. These studies did not specifically describe outcomes in the subgroup of children who would be identified by UNHS but not by selective screening.
3. What are the potential adverse effects of screening and early treatment? *Evidence is inconclusive.*	III	Poor: Most postulated adverse effects have not been evaluated in studies.

Note: Evidence codes: I: Randomized controlled trial; II-1: Controlled trial without randomization; II-2: Cohort or case-control analytic study; II-3: Multiple time series, dramatic uncontrolled experiments; III: Opinions of respected authorities

ABR indicates automated brainstem response; OAE, otoacoustic emissions; PHL, permanent hearing loss; UNHS, universal newborn hearing screening.

those who have had hearing aids and other therapy for shorter periods of time. None of these studies compared an inception cohort of newborns offered UNHS to infants managed by usual care (including selective screening). While they are better than older studies, these studies had unclear criteria for selecting subjects, making it impossible to exclude selection bias as an explanation for the findings. The hypothesis that early intervention is a predictor of language acquisition is plausible, but the studies do not establish that screening low-risk newborns is the important factor.

As use of UNHS rapidly expands, it is important to conduct longitudinal studies of UNHS to address these gaps in its scientific basis. Further randomized trials of UNHS seem unlikely to be conducted in the United States. Although it would be possible to compare states with and without UNHS, such studies would be prone to uncontrollable confounding due to differences among states. However, better evidence about the effectiveness of UNHS is needed and could be obtained via time-series, population-based studies that begin with inception cohorts and carefully report outcomes in all possible patients, 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. States that have UNHS should conduct such population-based studies to evaluate whether the long-term language outcomes of deaf children improve as the age of identification decreases.

References

1. Maki-Torkko EM, Lindholm PK, Vayrynen MRH, Leisti JT, Sorri MJ. Epidemiology of moderate to profound childhood hearing impairments in northern Finland. Any changes in ten years? *Scand Audiol.* 1998;27(2):95-103.
2. Fortnum H, Davis A. Epidemiology of permanent childhood hearing impairment in Trent Region, 1985-1993. *Br J Audiol.* 1997;31(6):409-446.
3. Karikoski JO, Marttila TI. Prevalence of childhood hearing impairment in southern Finland. *Scand Audiol.* 1995;24(4):237-241.
4. Parving A, Jensen JH. Prevalence of permanent childhood hearing impairment—its role in audit of local paediatric hearing health services. *J Audiol Med.* 1998;7(2):100-108.
5. Boyle CA, Yeargin-Allsopp M, Doernberg NS, Holmgreen MS, Murphy CC, Schendel DE. Prevalence of selected developmental disabilities in children 3-10 years of age: the metropolitan Atlanta developmental disabilities surveillance program, 1991. *MMWR, CDC Surv Summ.* 1996;45(SS-2):1-14.
6. Van Naarden K, Decoufle P, Caldwell K. Prevalence and characteristics of children with serious hearing impairment in metropolitan Atlanta, 1991-1993. *Pediatrics.* 1999;103(3):570-575.
7. Stein LK. Factors influencing the efficacy of universal newborn hearing screening. *Pediatr Clin North Am.* 1999;46(1):95-105.
8. Ruben RJ. Effectiveness and efficacy of early detection of hearing impairment in children. *Acta Otolaryngol.* 1991;482(Suppl):127-131.
9. Rach GH, Zielhuis GA, van den Broek P. The influence of chronic persistent otitis media with effusion on language development of 2- to 4- year-olds. *Int J Pediatr Otorhinolaryngol.* 1988;15:253-261.
10. Zinkus PW, Gottlieb MI. Patterns of perceptual and academic deficits related to early chronic otitis media. *Pediatrics.* 1980;66(2):246-253.
11. Moeller MP, Osberger MJ, Eccarius M. Receptive language skills. *Lang Learn Skills Hearing-Impaired Child.* 1986;23:41-53.
12. Carney AE, Moeller MP. Treatment efficacy: hearing loss in children. *J Speech Lang Hear Res.* 1998;41 (1):S61-84.
13. Yoshinaga-Itano C, Apuzzo ML. The development of deaf and hard of hearing children identified early through the high-risk registry. *Am Ann Deaf.* 1998;143(5):416-424.
14. Holt JA. Stanford Achievement Test-8th edition: reading comprehension subgroup results. *Am Ann Deaf Ref Iss.* 1993;138:172-175.
15. Allen TE. Patterns of academic achievement among hearing impaired students: 1974 and 1983. In: Schildroth A, Karchmer AM, eds. *Deaf Children in America.* Boston, MA: College-Hill Press; 1986:161-206.
16. National Institutes of Health. *Early Identification of Hearing Impairment in Infants and Younger Children.* Rockville, MD; 1993.

17. Fonseca S, Forsyth H, Grigor J, et al. Identification of permanent hearing loss in children: are the targets for outcome measures attainable? *Br J Audiol.* 1999;33(3):135-143.

18. Davis A, Bamford J, Wilson I, Ramkalawan T, Forshaw M, Wright S. A critical review of the role of neonatal hearing screening in the detection of congenital hearing impairment. *Health Technol Assess.* 1997;1(10):i-iv, 1-176.

19. Meadow-Orlans KP, Mertens DM, Sass-Lehrer MA, Scott-Olson K. Support services for parents and their children who are deaf or hard of hearing: a national survey. *Am Ann Deaf.* 1997;142(4):278-293.

20. Harrison M, Roush J. Age of suspicion, identification, and intervention for infants and young children with hearing loss: a national study. *Ear Hear.* 1996;17(1):55-62.

21. Kuhl PK, Williams KA, Lacerda F, Stephens KN, Lindbloom B. Linguistic experience alters phonetics perception in infants by six months of age. *Science.* 1992;255:606-608.

22. Gopnik A, Meltzoff AN, Kuhl PK. *The Scientist in the Crib: Minds, Brains, and How Children Learn.* New York: William Morrow and Company; 1999.

23. Sininger YS, Doyle KJ, Moore JK. The case for early identification of hearing loss in children. Auditory system development, experimental auditory deprivation, and development of speech perception and hearing. *Pediatr Clin North Am.* 1999;46(1):1-14.

24. Vohr BR, Carty LM, Moore PE, Letourneau K. The Rhode Island Hearing Assessment Program: experience with statewide hearing screening (1993-1996). *J Pediatr.* 1998;133(3):353-357.

25. Finitzo T, Diefendorf AO. The state of the information... evidence gathering in infant hearing programs. *Am J Audiol.* 1997;6(3):91-94.

26. White KR, Vohr BR, Behrens TR. Universal newborn hearing screening using transient evoked otoacoustic emissions: results of the Rhode Island Hearing Assessment Project. *Semin Hear.* 1993;14(1):18-29.

27. Joint Committee on Infant Hearing. Joint Committee on Infant Hearing Year 2000 Position Statement: Principles and Guidelines for Early Hearing Detection and Intervention Programs. *Pediatrics.* 2000;106(4):798-817. Available at: http://www.infanthearing.org/jcih/. Accessed October 15, 2001.

28. U.S. Preventive Services Task Force. Screening for hearing impairment. In: *US Preventive Services Task Force. Guide to clinical preventive services.* 2nd ed. Baltimore: Williams & Wilkins; 1996:393-405.

29. Grandori F, Lutman ME. European Consensus Statement on Neonatal Hearing Screening. Finalised at the European Consensus Development Conference on Neonatal Hearing Screening, 15-16 May 1998, Milan. *Int J Pediatr Otorhinolaryngol.* 1998;44:309-310.

30. Grandori F, Lutman M. The European Consensus Development Conference on Neonatal Hearing Screening (Milan, May 15-16, 1998). *Am J Audiol.* 1999;8:19-20.

31. Brookhouser PE. Sensorineural hearing loss in children. [Review]. *Pediatr Clin North Am.* 1996;43(6):1195-1216.

32. Eilers RE, Berlin C. Advances in early detection of hearing loss in infants. *Curr Probl Pediatr.* 1995;25(2):60-66.

33. Finitzo T, Crumley WG. The role of the pediatrician in hearing loss. From detection to connection. [Review]. *Pediatr Clin North Am.* 1999;46(1):ix-x, 15-34.

34. Martin WH, Schwegler JW, Gleeson AL, Shi YB. New techniques of hearing assessment. [Review]. *Otolaryngol Clin North Am.* 1994;27(3):487-510.

35. Oudesluys-Murphy AM, van Straaten HL, Bholasingh R, van Zanten GA. Neonatal hearing screening. [Review]. *Eur J Pediatr.* 1996;155(6):429-435.

36. Silverman CA. Audiologic assessment and amplification. *Prim Care.* 1998;25(3):545-581.

37. Erenberg A, Lemons J, Sia C, Trunkel D, Ziring P. Newborn and infant hearing loss: detection and intervention. American Academy of Pediatrics. Task Force on Newborn and Infant Hearing, 1998-1999. *Pediatrics.* 1999;103(2):527-530.

38. Wessex Universal Neonatal Hearing Screening Trial Group. Controlled trial of universal neonatal screening for early identification of permanent childhood hearing impairment. *Lancet.* 1998;352:1957-1964.

39. Mason JA, Herrmann KR. Universal infant hearing screening by automated auditory brainstem response measurement. *Pediatrics.* 1998;101:221-228.

40. Mehl A, Thomson V. Newborn hearing screening: the great omission. *Pediatrics.* 1998;E4:101.

41. Finitzo T, Albright K, O'Neal J. The newborn with hearing loss: dectection in the nursery. *Pediatrics.* 1998;102(6):1452-1460.

42. Prieve B, Dalzell L, Berg A, Bradley M, Cacace A. The New York State universal newborn hearing screening demonstration project: outpatient outcome measures. *Ear Hear.* 2000;21(2):131-140.

43. Barsky-Firkser L, Sun S. Universal newborn hearing screenings: a three-year experience. *Pediatrics.* 1997;99(6):E4.

44. Watkin PM. Neonatal otoacoustic emission screening and the identification of deafness. *Arch Dis Child Fetal Neonatal Ed.* 1996;74(1):F16-25.

45. Aidan D, Avan P, Bonfils P. Auditory screening in neonates by means of transient evoked otoacoustic emissions: a report of 2,842 recordings. *Ann Otol Rhinol Laryngol.* 1999;108(6):525-531.

46. Clemens CJ, Davis SA, Bailey AR. The "false positive" in universal newborn hearing screening. *Pediatrics.* 2000;106(1):7-11.

47. Norton SJ, Gorga MP, Widen JE, et al. Identification of neonatal hearing impairment: evaluation of transient evoked otoacoustic emission, distortion product otoacoustic emission, and auditory brain stem response test performance. *Ear Hear.* 2000;21(5):508-528.

48. Apuzzo ML, Yoshinaga-Itano C. Early identification of infants with significant hearing loss and the Minnesota child development inventory. *Semin Hear.* 1995;16(2):124-139.

49. Yoshinaga-Itano C, Sedey AL, Coulter DK, Mehl AL. Language of early- and later-identified children with hearing loss. *Pediatrics.* 1998;102(5):1161-1171.

50. Yoshinaga-Itano C, Apuzzo ML. Identification of hearing loss after age 18 months is not early enough. *Am Ann Deaf.* 1998;143(5):380-387.

51. Yoshinaga-Itano C, Coulter D, Thomson V. The Colorado Newborn Hearing Screening Project: effects on speech and language development for children with hearing loss. *J Perinatol.* 2000;20(8 Pt 2):S132-137.

52. Mayne AM, Yoshinaga-Itano C, Sedey AL, Carey A. Expressive vocabulary development of infants and toddlers who are deaf or hard of hearing. *Volta Review.* 2000;100(5):1-28.

53. Moeller MP. Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics.* 2000;106(3):E43-51.

54. Calderon R, Naidu S. Further support of the benefits of early identification and intervention with children with hearing loss. *Volta Review.* 2000;100(5):53-84.

55. Harris R, Helfand M, Woolf S, et al. Current Methods of the U.S. Preventive Services Task Force. *Am J Prev Med.* 2001;20(Suppl 3):21-35.

56. Dalzell L, Orlando M, MacDonald M, et al. The New York State universal newborn hearing screening demonstration project: ages of hearing loss identification, hearing aid fitting, and enrollment in early intervention. *Ear Hear.* 2000;21(2):118-130.

57. Kennedy CR. Controlled trial of universal neonatal screening for early identification of permanent childhood hearing impairment: coverage, positive predictive value, effect on mothers and incremental yield. Wessex Universal Neonatal Screening Trial Group. *Acta Paediatr Suppl.* 1999;88(432):73-75.

58. Folsom RC, Diefendorf AO. Physiologic and behavioral approaches to pediatric hearing assessment. *Pediatr Clin North Am.* 1999;46(1):107-120.

59. Johnson JL, Kuntz NL, Sia CC, White KR, Johnson RL. Newborn hearing screening in Hawaii. *Hawaii Med J.* 1997;56(12):352-355.

60. Prieve BA, Stevens F. The New York State universal newborn hearing screening demonstration project: introduction and overview. *Ear Hear.* 2000;21(2):85-91.

61. Thomson V. The Colorado Newborn Hearing Screening Project. *Am J Audiol.* 1997;6(3):74-77.

62. Thomson V, Rose LB, O'Neal J, Finitzo T. Statewide implementation of universal newborn hearing screening. *Semin Hear.* 1998;19(3):287-300.

63. Watkin P, Baldwin M, Dixon, R, Beckman A. Maternal anxiety and attitudes to universal neonatal hearing screening. *Br J Audiol.* 1998:27-37.

64. Barringer DG, Mauk GW. Survey of parents' perceptions regarding hospital-based newborn hearing screening. *Audiol Today.* 1997;9:18-19.

65. Priesler G. The development of communication and language in deaf and severely hard of hearing children: implications for the future. *Int J Pediatr Otorhinolaryngol.* 1999;49(Suppl 1):S39-43.

Appendix 1. Search strategy

Set	Search
1	exp hearing disorders/
2	infant/or infant, newborn/
3	1 and 2
4	limit 3 to human/
5	limit 4 to English language/
6	4 not 5
7	limit 6 to abstracts
8	5 or 7
9	exp mass screening/
10	screen$.tw.
11	exp hearing tests/
12	9 or 10 or 11
13	8 and 12
14	cochlear implants/
15	exp hearing aids/
16	exp manual communication/
17	exp rehabilitation of hearing impaired/
18	esp hearing disorders/dt,rh,su,th
19	14 or 15 or 16 or 17 or 18
20	8 and 12
21	13 or 20
22	exp hearing disorders/
23	limit 22 to human
24	limit 23 to English language
25	limit 24 to (preschool child <2 to 5 years> or child <6 to 12 years> or adolescence <13 to 18 years>)
26	19 and 25
27	exp evaluation studies/
28	follow-up studies/
29	meta analysis/
30	exp clinical trials/
31	27 or 28 or 29 or 30
32	26 and 31
33	limit 26 to (controlled clinical trial or guideline or meta analysis or mulitcenter study or practice guideline or randomized controlled trial or review, multicase)
34	32 or 33
35	21 or 34

Appendix 2. U.S. Preventive Services Task Force Quality Rating

Randomized Controlled Trials and Cohort Studies

Criteria:

- Initial assembly of comparable groups.
  a. For RCTs: adequate randomization, including first concealment and whether potential confounders were distributed equally among groups.
  b. For cohort studies: consideration of potential confounders with either restriction or measurement for adjustment in the analysis; consideration of inception cohorts.
- Maintenance of comparable groups (includes attrition, cross-overs, adherence, contamination).
- Levels of follow-up: differential loss between groups; overall loss to follow-up
- Measurements: equal, reliable, and valid, and including masking of outcome assessment.
- Clear definition of interventions.
- Important outcomes considered.
- Analysis:
  a. For RCTs: intention-to-treat analysis
  b. For cohort studies: adjustment for potential confounders.

Definition of ratings based on above criteria:

Good: Meets all criteria: comparable groups are assembled initially and maintained throughout the study; follow-up at least 80 percent; reliable and valid measurement instruments applied equally to the groups; interventions clearly defined; important outcomes are considered; and appropriate attention to confounders in analysis. In addition, for RCTs, intention-to-treat analysis is used.

Fair: Generally comparable groups assembled initially but some question remains whether some (although not major) differences occurred in follow-up; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for. Intention-to-treat analysis is done for RCTS.

Poor: Groups assembled initially are not close to being comparable or maintained throughout the study; measurement instruments are unreliable or invalid or not applied at all equally among groups; outcome assessment not masked; and key confounders are given little or no attention. For RCTs, no intention-to-treat analysis.

AHRQ Pub. No. 02-501B
October 2001