From the Department of Clinical Science, Intervention and Technology,
Division of Ear, Nose and Throat Diseases,
Karolinska Institutet, Stockholm, Sweden

PERSPECTIVES ON SCREENING STRATEGIES FOR EARLY DETECTION OF CHILDHOOD HEARING IMPAIRMENT

Allison Ruth Mackey

Stockholm 2022

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2022

ISBN 978-91-8016-729-1
Cover illustration by Gustav Hagerling

Perspectives on screening strategies for early detection of childhood hearing impairment

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Allison Mackey

The thesis will be defended in public at B64 Karolinska University Hospital Huddinge, Stockholm Sweden, on the 14th of October 2022.

Principal Supervisor:
Associate Professor Inger Uhlén
Karolinska Institutet
Department of Clinical Science,
Intervention and Technology
Division of Ear, Nose and Throat Diseases

Co-supervisor:
Professor Elina Mäki-Torkko
Örebro University
Faculty of Medicine and Health
Audiological Research Centre

Opponent:
Professor Elizabeth Fitzpatrick
University of Ottawa
Faculty of Health Sciences
School of Rehabilitation Sciences

Examination Board:
Associate Professor Jonas Brännström
Lund University
Department of Clinical Sciences
Division of Logopedics, Phoniatrics and Audiology

Professor Ona Bø Wie
University of Oslo
Department of Department of Special Needs Education

Associate Professor Birger Forsberg
Karolinska Institutet
Department of Global Public Health

POPULAR SCIENCE SUMMARY OF THE THESIS

One of every 500 newborns have a permanent hearing loss. A permanent hearing loss is a life-long condition. Infants learn language by listening to others around them. With exposure to language during the development of the nervous system, the brain learns to decode auditory signals into meaningful components, such as words and sentences. Language sets the stage for reading, social communication, and education. This nervous system development happens quickly within the first few months of life and slows down over the lifetime. This is why it is vital that intervention is provided to infants and children with hearing loss as early as possible. Missing the optimal window for language development can have lifetime consequences.

It is very difficult to notice a hearing loss, particularly among newborns. In the last few decades, researchers have developed tools that can quickly and automatically measure responses to sound from the ear and auditory pathway. These tools are called the otoacoustic emissions (OAE) test and the automated auditory brainstem response (aABR) test. Newborn hearing screening uses this technology. Screening ends in a pass or fail result. Newborns who fail may be screened a second time a few days or weeks later. If they still fail, they should be tested more thoroughly in a diagnostic clinic. Hearing tests in in older children use behavioral methods, in which the child responds to each sound by performing an action, such as moving a toy. These methods are used at health checkups around the age of 3 to 7 years, while children are in preschool or entering school.

Screening programmes are complex. They take careful planning to organize, implement and maintain. A failed hearing screen does not necessarily mean that the baby has a hearing loss. Some may have hearing loss, but most babies who fail have normal hearing. They fail screening for other reasons. For example, fluid in the ear that hasn't yet drained out since birth, or too much background noise, can cause a screening fail. Too many fails are costly for a programme, put too much pressure on diagnostic clinics, and cause unnecessary worry for families. Conversely, programmes are not effective if too many children with hearing loss pass screening. Many countries have implemented hearing screening programmes. How well are existing programmes performing when it comes to screening for and detecting hearing loss? How can programmes plan or revise their protocol to improve their outcomes?

There are also different setups for a screening programme. For example, are all newborns or only risk groups screened? Who performs the screening and where? When is screening performed, just hours after birth or a few days later? What test method is used for screening? For newborns who failed the screen, are they rescreened a few days to weeks later before they are sent for diagnostic testing? If so, how many rescreening appointments are planned?

The purpose of this thesis was to evaluate hearing screening programmes. Three different methods were used across five studies. In **studies I, II and III**, as part of the EU project EUSCREEN, professionals across 47 countries or regions who represented hearing screening programmes were surveyed. In **study IV**, all scientific studies that evaluated one of two

outcomes from screening were reviewed: the percentage of infants that were referred from screening or the percentage of infants that returned for follow-up testing after a failed result. Finally, **in study V**, the efficacy of the passing criteria used in the newborn screening program in Region Stockholm was studied.

Studies I, II and III

The survey used in **studies I, II and III** included questions about the status of the hearing screening programme. Results of **study I** showed that most high-income countries performed newborn hearing screening on all infants, however, only a few middle-income countries had similar programmes.

Study III showed a different trend for preschool- and school-entry screening. Only 17 out of 47 countries or regions had programmes that screened all children at preschool or school-entry age. Some countries have started relatively new screening programmes. At the same time, other countries have scaled back or stopped their programmes.

Study I showed that in most newborn hearing screening programmes nurses or midwives performed the screening. Other professionals who performed screening were audiologists or physicians. Some programmes hired individuals specifically for the task of hearing screening. Most screening was performed in the maternity ward before the baby and mother were sent home. For a few programmes, newborns were screened later, at home or in well-baby clinics. Most programmes screened newborns 24 to 72 hours after birth, though a few screened within 24 hours and a few screened after 72 hours.

OAE was most commonly used for screening well babies, likely for reasons of simplicity and lower costs. Most programmes used aABR for babies in the neonatal intensive care unit. This is recommended because babies that need intensive care have a higher chance of hearing loss located in the auditory nerve, which can be detected with aABR. Some programmes also used aABR for well babies, typically during a rescreening appointment. Most commonly, newborns who failed were rescreened one or two times before being referred for diagnostic testing. Countries that used aABR and multiple rescreening appointments had higher health care spending per person, on average. On the other hand, some programmes used OAE for all babies, including those who needed intensive care. These countries had, on average, lower health care spending per person.

Study III showed that, for preschool- or school-entry hearing screening, children were screened between ages 3 and 7, either in a health clinic or at the school or preschool. The screening intensity that determined a passing result ranged from 20 dB to 40 dB HL across programmes. Nurses most commonly performed screening.

Studies II and III evaluated the quality of screening programmes. Questions included: what percentage of children are screened? What percentage fail screening? What percentage of children who failed screening attend their follow-up appointment? And what percentage of children screened are found to have hearing loss?

Results from **studies II and III** showed that few respondents provided answers to these questions based on actual data. For example, less than one third could provide a high-quality answer about the percentage of newborns that attended their follow-up appointment after referring from screening. Similarly, these data were only available for two out of 17 preschool or school-entry programmes.

Study II showed a lower percentage of babies who were referred from screening for programmes that used aABR for well babies instead of OAE and for programmes that rescreened newborns two or more times before referring them to the diagnostic clinic. This study also showed that the follow-up rate from screening to diagnostic assessment was less than 90% for eight out of the 12 programmes that provided this outcome.

Study IV

Study IV looked further into these issues: What are the factors in a screening programme that can reduce the percentage of infants who fail, and the percentage who are lost to follow-up? A systematic review of the scientific literature was performed.

There were some key factors that were related to the percentage of infants who failed screening. To reduce the percentage of fails, programme decision-makers might consider screening babies around 3 to 5 days of age, using aABR instead of OAE, try screening again after an initial failed result, wait for babies to settle before screening, and network smaller screening sites to larger centers to increase the capacity for support and quality management. Anyone making decisions regarding their screening programme should first consider the needs and circumstances in their screening programme before making these adjustments, to identify whether these parameters would be effective in their particular context.

High loss to follow-up is a concerning outcome. Babies who failed screening who do not return for follow-up may have a hearing loss that remains undetected. According to some studies, infants who had a risk factor for hearing loss were more likely to be lost to follow-up than infants with no risk factor. This is particularly concerning.

To improve loss to follow-up, decision-makers might consider using a familiar and accessible location for the follow-up appointment, schedule the follow-up appointment in person with the family, and ensure the sufficiency of workload burden, experience, and knowledge of the professionals involved in the screening programme. One study showed that the involvement of an audiologist in the screening programme was the most significant factor in achieving good follow-up rates (Thomson & Yoshinaga-Itano, 2018). However, overall, there were few high-quality studies on loss to follow-up.

Study V

Some infants may not be detected by newborn hearing screening because the technology used for screening indicated a pass when the newborn actually had a hearing loss. Pre-decided criteria in the OAE and aABR equipment will determine a pass versus a fail result. Newborns

who pass hearing screening despite having a hearing loss may not be detected until many years later. Eventually, they may be detected by other methods, such as caregiver concern or preschool hearing screening. All children in Region Stockholm who are receiving intervention for their hearing loss are entered into a registry. The newborn hearing screening results were studied for children with hearing loss who were enrolled for intervention before 8 years of age. For the children who were documented as having passed screening, what did the results of their OAEs show?

Study V showed two important results. First, 11% of children who reportedly passed OAE should not have passed according to their OAE data. This was likely because of an error by the screener who entered in a pass by mistake. Second, two OAE measures were shown to significantly predict the group of ears with hearing loss, compared to a group of ears with normal hearing. This suggests that, if passing criteria were stricter, then more children with hearing loss would have been detected at birth.

Conclusions

(1) Progress is needed so that newborn hearing screening exists in all countries, and all children with hearing loss are given the opportunity for early intervention. **(2)** There is no consensus across countries regarding the importance of hearing screening at preschool or school-entry age. **(3)** Lost to follow-up is a problem faced by many hearing screening programmes. **(4)** Some factors may reduce the percentage of babies who refer from screening. These factors include: the screening method (aABR instead of OAE), the number of times screening is reattempted, the age of the babies when they are screened, the experience of the person performing the screening, and how the screening site is integrated with other sites. **(5)** Some factors may reduce the percentage of babies who are lost to follow-up. These factors include the location of the follow-up appointment, the experience and workload of the screeners, and whether an audiologist is involved in the programme. **(6)** If passing criteria are stricter, more babies with hearing loss will be detected by newborn hearing screening. **(7)** Finally, improvements are needed in data collection, reporting, and monitoring of outcomes. How do programmes identify areas for improvement if outcomes are not collected or reported? The first step toward improving the quality of a screening programme is to set up a system for systematically monitoring and evaluating the outcomes from screening. Once that is established, the other findings from this thesis can be applied.

ABSTRACT

The aims of this thesis were to assess the status of childhood hearing screening programmes, primarily in Europe, and to evaluate their performance against the parameters that make up a childhood hearing screening programme, such as test method, protocol, location, screening professional, and age of the infant at screening.

Studies I, III and III were ecological studies within the EUSCREEN project. The purpose of the EUSCREEN project was to develop a cost-effectiveness model and toolkit for implementation and modification of childhood hearing and vision screening programmes. Studies I, II and III made up an independent line of research that assessed existing childhood hearing screening programmes. A comprehensive questionnaire was delivered to professionals representing their local hearing screening programme. Questions included the provision, protocols, factors, and performance of newborn and childhood hearing screening. **Study IV** was a systematic review of literature. Studies were aggregated that compared referral or follow-up rate between parameters in a newborn hearing screening (NHS) programme. **Study V** investigated the presence of false negatives in the NHS programme in Region Stockholm. A retrospective analysis of the otoacoustic emission (OAE) results was performed among children with hearing impairment who had previously passed NHS.

Universal nationwide NHS existed in 25 of 30 high-income countries surveyed and 3 of 15 low- to middle-income countries. Universal preschool or school-entry screening existed in 17 out of 47 countries or regions. For NHS, countries that only used OAE for the test method had lower health spending compared to those that used automated auditory brainstem response (aABR). However, using aABR resulted in lower referral rates compared to OAE. Other factors that influenced referral rate were the number of rescreens, the age of the infant, the experience of the screening professional and the organization of the screening programme.

Out of the 12 programmes with a valid follow-up rate to diagnostic assessment, eight had rates below 90%. Factors such as personnel experience and knowledge and the location of follow-up can improve rates; however, few high-quality studies investigated this issue. There are other reasons why children with hearing impairment are not detected by NHS. Out of 1244 children with hearing impairment in study V, 24 were lost due to errors in documentation. Additionally, two OAE variables predicted hearing impairment among infants who previously met passing criteria ($p<0.001$). Increasing the stringency of passing criteria will increase the number of children with hearing impairment detected by screening.

The findings from this thesis may be used for implementing new programmes or improving existing ones. However, the results also suggest a significant deficiency in the process for monitoring and evaluation of hearing screening programmes. Out of 42 NHS programmes surveyed, 23 had valid referral rates and 12 had a valid follow-up rate. Out of 17 preschool or school-entry programmes, only two provided these data. Developing a process for collecting and reporting on existing outcomes is the necessary first step to quality improvement.

LIST OF SCIENTIFIC PAPERS

I. Bussé AML†, **Mackey AR†**, Hoeve HLJ, Goedegebure A, Carr G, Uhlén IM, et al. Assessment of hearing screening programmes across 47 countries or regions I: provision of newborn hearing screening. International Journal of Audiology. 2021;60(11):821-830.

II. **Mackey AR†**, Bussé AML†, Hoeve HLJ, Goedegebure A, Carr G, Simonsz HJ, et al. Assessment of hearing screening programmes across 47 countries or regions II: coverage, referral, follow-up and detection rates from newborn hearing screening. International Journal of Audiology. 2021;60(11):831-840.

III. Bussé AML†, **Mackey AR†**, Carr G, Hoeve HLJ, Uhlén IM, Goedegebure A, et al. Assessment of hearing screening programmes across 47 countries or regions III: provision of childhood hearing screening after the newborn period. International Journal of Audiology. 2021;60(11):841-848.

IV. **Mackey AR,** Bussé AML, Del Vecchio V, Mäki-Torkko E, Uhlén IM. Protocol and programme factors associated with referral and loss to follow-up from newborn hearing screening: a systematic review. BMC Pediatrics. 2022;22:473.

V. **Mackey AR**, Mäki-Torkko E, Uhlén, IM. Hearing loss missed by newborn hearing screening: Revisiting transient-evoked otoacoustic emissions after a passing result. [*Submitted*]

(† Dual first authorship)

CONTENTS

LIST OF ABBREVIATIONS

aABR	Automated auditory brainstem response
ABR	Auditory brainstem response
ANSD	Auditory neuropathy spectrum disorder
BOEL	Blicken Orienterar efter Ljudet
CSOM	Chronic suppurative otitis media
dB HL	Decibel hearing level
dB nHL	Decibel normalized hearing level
DPOAE	Distortion product otoacoustic emissions
EEG	Electroencephalogram
EHDI	Early hearing detection and intervention
EU	European Union
HI	Hearing impairment
LMIC	Low- and middle-income countries
LTFU	Lost to follow-up
NHS	Newborn hearing screening
NICU	Neonatal intensive care unit
OAE	Otoacoustic emissions
OME	Otitis media with effusion
SNR	Signal to noise ratio
TEOAE	Transient-evoked otoacoustic emissions

1 INTRODUCTION

Hearing plays the essential role in stimulating the development of central auditory pathways in early childhood. Approximately 1 to 3 per 1000 infants are born with a permanent hearing impairment (HI) (Bussé et al., 2020; Butcher et al., 2019; Watkin & Baldwin, 1999). HI can have permanent effects on language (Blamey et al., 2001; Ching et al., 2010; Moeller, 2000; Yoshinaga-Itano et al., 1998), literacy (Moeller et al., 2007; Tomblin et al., 2020), and psychosocial development (Vernon, 1969; Wake et al., 2004; Wong et al., 2017). For this reason, early intervention is critical.

Newborn hearing screening (NHS) has changed the scope of practice for early hearing detection and intervention (EHDI). With NHS, HI can be detected, diagnosed and intervened within weeks to months after birth (Uus & Bamford, 2006; Yoshinaga-Itano, 2003). The Joint Committee on Infant Hearing, in their so-called '1-3-6' framework, set benchmarks for diagnosis, detection and intervention at 1-month, 3-month and 6-months of age (Joint Committee on Infant Hearing, 2007, 2019). The World Health Organization recommends that countries and regions around the world strive for universal NHS. If resources are insufficient, then a non-universal (e.g., targeted) programme is a positive first step toward scaling up to widespread universal NHS (*Hearing screening: considerations for implementation*, 2021; Tobe et al., 2013; World Health Organization, 2021). With targeted screening, only infants with risk factors for HI are screened, and about 50% of newborns with HI are detected (Mauk et al., 1991; Mehl & Thomson, 1998).

Even with universal NHS, not all children with HI are detected. Delayed-onset and acquired HI contributes to an increasing prevalence of HI throughout early childhood (Lü et al., 2011; Uhlén et al., 2020; Watkin & Baldwin, 2012). Furthermore, it is possible that some children with HI pass NHS (Norton, Gorga, Widen, Folsom, et al., 2000). Childhood hearing screening after the newborn period (i.e., preschool or school-entry screening) pre-exists NHS (Ewing, 1955; Fisch, 1981; Reznik et al., 1985). Since implementation of NHS, the necessity of preschool and school-entry screening has been questioned (Fortnum et al., 2016).

There is a diversity of parameters and protocols used for childhood hearing screening (Sloot et al., 2015). The components of a programme include the screening method, the passing criteria, the number of rescreening steps, who performs the screening, when screening is performed, the location where screening is performed, and who manages, funds, and governs the programme. How do decision-makers select the most appropriate set of parameters for their screening programme? The sensitivity and specificity are typical measures used to describe the effectiveness of a screening programme; however, they are not easily quantified for childhood hearing screening. Instead, quality indicators are used. These are predefined and measurable outcomes. For childhood hearing screening, quality indicators include the coverage rate, referral rate, follow-up rate, and detection rate (Joint Committee on Infant Hearing, 2019). The purpose of this thesis was to map the provision of childhood hearing screening and evaluate the parameters, protocols and outcomes of existing programmes.

2 LITERATURE REVIEW

2.1 CHILDHOOD HEARING IMPAIRMENT

Permanent HI affects one to three per 1000 infants (up to age 1 year) (Bussé et al., 2020; Butcher et al., 2019; Watkin & Baldwin, 1999). The prevalence rate increases by a factor of three by 7 years of age (Uhlén et al., 2020). The following sections describe the definition and prevalence of childhood HI.

2.1.1 Definition

In this thesis, childhood HI is defined based on hearing threshold levels. Seven domains that describe childhood HI are presented in this thesis: onset, longevity, stability, type, laterality, configuration, and degree. These are illustrated in Figure 1.

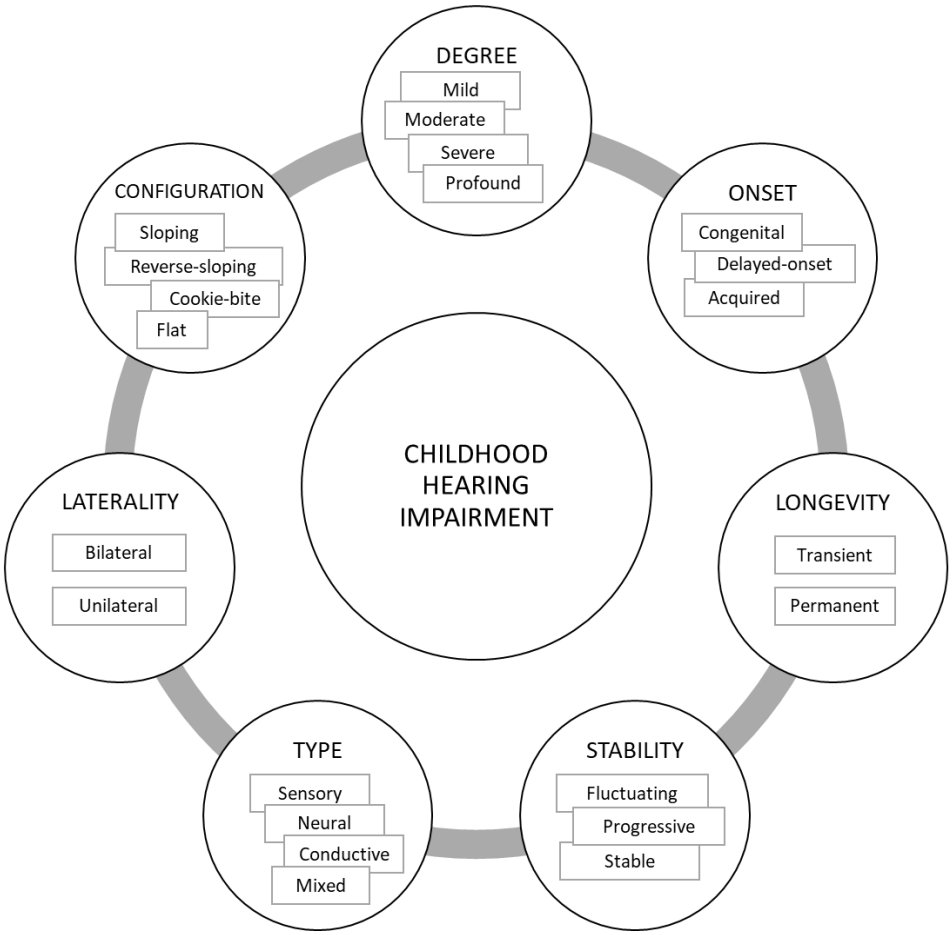

Figure 1. Domains of childhood hearing impairment

Onset explains when the HI developed. Congenital HI is present at birth. Delayed onset defines a HI that manifests after birth though the conditions for HI were present during or shortly after birth. Acquired HI developed after birth from extrinsic factors unrelated to the birth.

HI can be classified as neural, sensory, conductive or mixed, based on its source in the auditory pathway. Conductive HI is localized to the middle or outer ear. A sensorineural HI is localized in the cochlea and/or neural pathways. Sensory HI is localized to the hair cells of the cochlea, while a neural HI is localized to the auditory nerve or brainstem. An example of a neural HI among children is auditory neuropathy spectrum disorder (ANSD). For children with ANSD, the outer and middle ears function normally, as do the outer hair cells of the inner ear. The HI is caused by a transmission breakdown along the auditory nerve, from the inner hair cell to the brainstem (Starr et al., 1996). A mixed HI is a combination of both a conductive and sensorineural component.

Childhood HI can be transient or permanent. With permanent HI, thresholds cannot return to normal levels. The majority of permanent HI in high-income countries is sensorineural (Fortnum & Davis, 1997). Permanent HI may be stable (the degree of HI remains constant over time) or progressive (worsening of hearing thresholds over time). Thresholds associated with a transient HI can return to normal levels in time or with intervention.

Other domains of HI include its degree, whether it exists in one or both ears (unilateral or bilateral) and the configuration of the hearing thresholds across frequencies. The configuration is the shape of the HI across frequencies. The degree is denoted in decibel hearing level (dB HL). The recent World Report on Hearing classifies degree using a pure-tone average of 0.5, 1, 2 and 4 kHz. Mild HI is between 20 to 34 dB HL, moderate HI is 35 to 49 dB HL, moderately severe HI is 50 to 64 dB HL, severe HI is 65 to 79 dB HL, profound HI is 80 to 94 dB HL, and complete HI is 95 dB HL or greater (World Health Organization, 2021). One definition of a unilateral HI is a pure-tone average of more than 20 dB HL for the frequencies 500, 1000 2000 Hz, or more than 50 dB HL at one frequency, while the opposite ear is normal (Davis & Davis, 2016).

2.1.2 Prevalence rates

Prevalence rates of permanent childhood HI vary across age and region (Bussé et al., 2020; Butcher et al., 2019; Uhlén et al., 2020), and between infant populations, e.g., well infants versus infants who were admitted to the neonatal intensive care unit (NICU) (Bussé et al., 2020; Butcher et al., 2019). For infants up to one year of age, the pooled prevalence of permanent congenital HI is 1.1 per 1000 for bilateral HI (Butcher et al., 2019) and 2.2 per 1000 for bilateral or unilateral HI (Bussé et al., 2020). Both reviews showed a significantly higher prevalence of permanent HI among infants who had been admitted to the NICU (Bussé et al., 2020; Butcher et al., 2019).

Studies from the U.K. (Fortnum et al., 2001), Australia (Ching et al., 2006) and Sweden (Uhlén et al., 2020) showed similar trends, that the prevalence of permanent HI significantly increases with age. The prevalence of permanent HI in Region Stockholm for ages 1, 5 and 7 years is described in Table 1, adapted from Uhlén et al. (2020). These data are based on the number of children receiving intervention for HI.

Table 1. Prevalence of permanent hearing impairment (HI) per 1000 children in Region Stockholm, adapted from Uhlén et al. (2020). Degree is calculated as a pure-tone average of 0.5, 1, 2 and 4 kHz.

Ear(s) and degree of HI	1-year olds	5-year olds	7-year olds
Bi- and unilateral permanent HI > 20 dB HL	1.2	2.7	3.5
Bilateral permanent HI >40 dB HL BE	0.7	1.3	1.7
Unilateral permanent HI >40 dB HL WE	0.3	0.7	0.9
Bilateral permanent HI 21-40 dB HL BE	0.2	0.7	0.9

BE: better ear; WE: worse ear; dB HL: decibel hearing level

The prevalence of ANSD is important to consider when evaluating the target condition for NHS and the technology used. The prevalence reported in the literature varies widely, in part due to the differences in methodology for diagnosing and defining HI and ANSD (Rance & Starr, 2017). In studies of children with permanent HI, the percentage of children with ANSD is approximately 5-10% (Boudewyns et al., 2016; Rance & Starr, 2017). Within a population of infants at risk (e.g., from the NICU), the reported prevalence rate ranges from 0.23% (Rance et al., 1999) to 9.6% (Berg et al., 2005). For populations of infants not at risk (e.g., well babies), the prevalence rate of ANSD ranges from 0.006 to 0.03% (Boudewyns et al., 2016; Korver et al., 2012). The term "risk" is used to denote infants with apparent risk factors. In this thesis, the term "low-risk" is also used to classify infants without risk factors, and thereby distinguish between infant groups and NHS protocols more easily.

2.2 EARLY HEARING DETECTION AND INTERVENTION

Advocacy for EHDI started in the 1960s to 1970s when the Joint Committee on Infant Hearing was formed (Committee on Fetus and Newborn, 1971). Significant momentum was gained in the 1990s and early 2000s in light of landmark studies on the effects of HI intervention on language development, if provided within a critical period (Moeller, 2000; Moeller et al., 2007; Yoshinaga-Itano et al., 1998). Further to this work, the effects of EHDI on literacy (Pimperton et al., 2014), psychosocial development and quality of life (Korver et al., 2010) have substantiated the importance of early detection and intervention.

2.2.1 Detecting neonatal hearing impairment

Detecting HI in newborns is the first step toward early intervention. A high-risk register was an early method for detecting HI (Hirsch & Kankkunen, 1974). It was recommended that infants with at least one risk factor were monitored and tested by 6 months of age (Joint Committee on Infant Hearing, 1982). Since the development of technology for testing infants, such as ABR and OAE, the high-risk register transitioned to targeted hearing screening in which only neonates in risk categories were screened (Joint Committee on Infant Hearing, 1991; Watson et al., 1996). However, only 50% of infants with HI are detected by targeted screening (Mauk et al., 1991; Mehl & Thomson, 1998). Studies showing the feasibility and effectiveness of universal NHS (Mehl & Thomson, 1998; Wessex Universal Neonatal Hearing Screening Trial Group, 1998) led to widespread advocacy for universal NHS (Joint

Committee on Infant Hearing, 1995, 2000, 2007, 2019). Benchmarks for the age of detection and intervention of congenital HI were published. NHS programmes should strive for the completion of screening by 1 month, diagnosis of HI by 3 months and intervention by 6 months of age (Joint Committee on Infant Hearing, 2007). Given the potential benefits of even earlier intervention (Walker et al., 2022), recently published guidelines recommend the provision of intervention before 3 months age where possible (Joint Committee on Infant Hearing, 2019).

2.2.2 Detecting hearing impairment after the newborn period

The increasing prevalence with age may be due to delayed-onset HI, acquired HI, or congenital HI that was not detected by NHS (Watkin & Baldwin, 2012). With delayed-onset or acquired HI, neonatal hearing was normal, and HI manifested in early childhood. Congenital HI that was not detected by NHS could be because the infant was never screened, was loss to follow-up after referral, or had a mild HI that was missed by NHS protocols (Watkin & Baldwin, 2012). Historically, the target condition for NHS is a bilateral HI of >40 dB HL (Wessex Universal Neonatal Hearing Screening Trial Group, 1998). A mild HI at birth may be progressive, leading to more severe degrees in early childhood (Barreira-Nielsen et al., 2016). Finally, infants with HI may be missed by NHS due to screener or algorithm error leading to a false negative. However, studies reporting on the false negatives from NHS are limited (Johnson et al., 2005; Norton, Gorga, Widen, Folsom, et al., 2000; Wessex Universal Neonatal Hearing Screening Trial Group, 1998).

Strategies for detecting HI beyond the newborn period include targeted surveillance of children with risk factors for delayed-onset HI, ongoing monitoring of development milestones, and universal screening in early childhood (Beswick, Driscoll, & Kei, 2012). With targeted surveillance, a registry of children with risk factors for delayed-onset HI is formed, and regular follow-ups and assessments are performed on these children at predetermined intervals. This method is effective at identifying some children with postnatal HI (Beswick, Driscoll, Kei, et al., 2012); however, 26 to 62% of children with HI identified after the newborn period do not present with risk factors (Lü et al., 2011; Weichbold et al., 2006). Furthermore, issues include loss to follow-up, discrepancy regarding the risk factors associated with postnatal HI, and the difficulty identifying all risk factors at birth (Beswick, Driscoll, & Kei, 2012; Beswick, Driscoll, Kei, et al., 2012).

Another method for detecting HI after the newborn period is by universal screening. Prior to NHS, two common ages for screening were around 7 to 10 months of age and around 3 to 7 years of age. Programmes that were previously in place in some countries for universally screening HI in infants ages 7-10 months of age became obsolete after implementation of NHS (Wood et al., 1997). In Sweden, this programme used the Blicken Orienterar efter Ljudet (BOEL) test, in which the examiner used small bells concealed in his hand and observed the reaction of the infant who was expected to turn or gaze toward to the direction of the sound (Junker et al., 1978). In England, a similar test with a slightly different procedure

was called the health visitor distraction test (Johnson & Ashurst, 1990; Mott & Emond, 1994). The replacement of the distraction test with NHS proved to be a cost-saving adjustment (Uus et al., 2006). Universal hearing screening at pre-school and school entry age (e.g., 3 to 7 years old) is recommended by some, including the World Health Organization (*Hearing screening: considerations for implementation*, 2021; Lü et al., 2011; Weichbold et al., 2006; World Health Organization, 2021). However, this recommendation is not universally accepted, as the cost-effectiveness of early childhood hearing screening is inconsistent across the few available studies (Fortnum et al., 2016; Gumbie et al., 2022).

2.3 SCREENING

Screening in health care is a strategy to uncover an existing disease or condition in an individual where it otherwise would not be recognized. Screening is classified as a secondary measure of prevention. Primary measures are to prevent the disease or condition from occurring, while secondary measures aim to reduce its impact. The purpose of screening is to identify conditions early, where earlier intervention provides a greater likelihood of success. The following sections define the principles of screening and the effectiveness of screening.

2.3.1 Principles of screening

To avoid implementation of unnecessary, harmful, or costly screening practices, ten principles of screening were established by the World Health Organization (Wilson & Jungner, 1968). Since the publication of these principles, Andermann et al. (2008) have proposed revisions to meet modern technological advances in screening and technology (Table 2). Any screening programme should fulfill all 10 requirements.

Table 2. Andermann et al.'s (2008) revised principles of screening. Reprinted with permission and under the Creative Commons Attribution 3.0 IGO license (CC BY 3.0 IG0).

1. The screening programme should respond to a recognized need.
2. The objectives of screening should be defined at the outset.
3. There should be a defined target population.
4. There should be scientific evidence of screening programme effectiveness.
5. The programme should integrate education, testing, clinical services and programme management.
6. There should be quality assurance, with mechanisms to minimize potential risks of screening.
7. The programme should ensure informed choice, confidentiality and respect for autonomy.
8. The programme should promote equity and access to screening for the entire target population.
9. Programme evaluation should be planned from the outset.
10. The overall benefits of screening should outweigh the harm.

2.3.2 Effectiveness of screening

Effectiveness is covered in the fourth principle proposed by Andermann et al. (2008). It is a general term to describe whether screening works in a field (real world) setting (Hyde, 2016). A decision matrix presented in Figure 2 is commonly used to describe the effectiveness of any screening or diagnostic test. Sensitivity is the percentage of individuals with the target condition who have a positive result. Specificity is the percentage of individuals without the target condition who have a negative result. The false negative rate is the percentage of individuals with the target condition with a negative result (1-sensitivity). The false positive rate is the percentage of individuals without the target condition with a positive result (1-specificity).

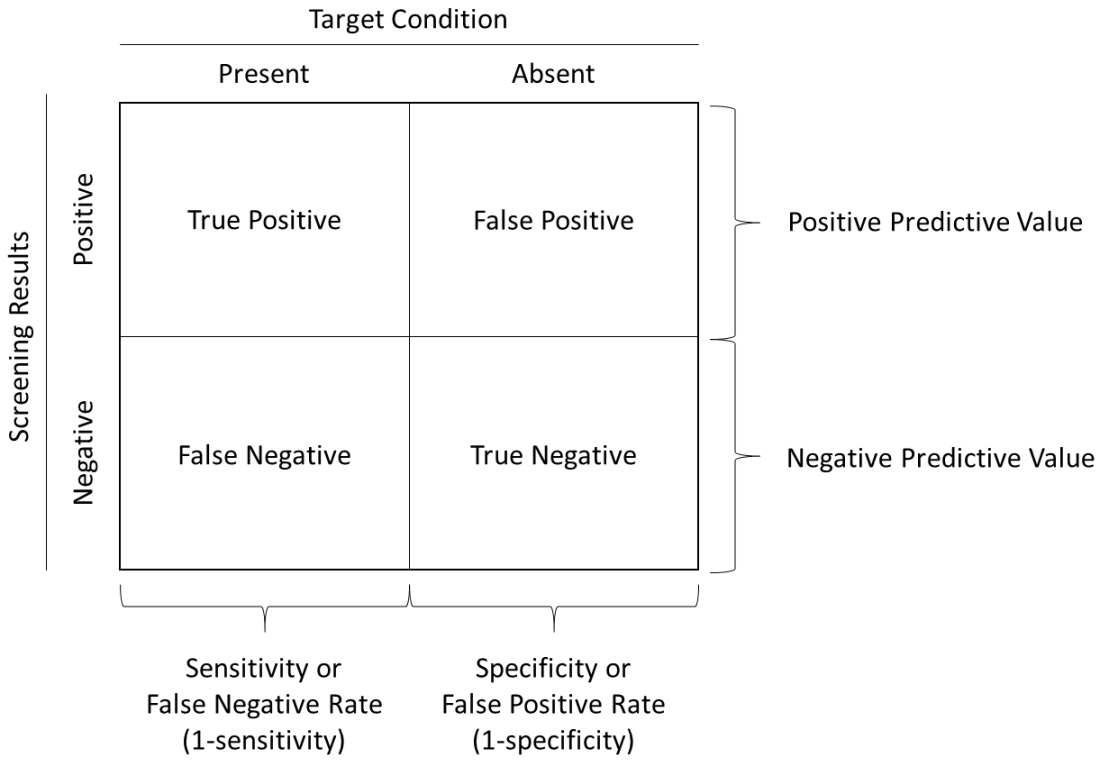

Figure 2. Decision matrix and measures of screening effectiveness.

2.3.3 Considerations on measuring the effectiveness of childhood hearing screening

It is challenging to definitively determine the sensitivity and specificity of NHS. Specificity can be reasonably estimated from the referral rate because of the relatively low prevalence rate of HI. Estimating the sensitivity is more challenging. To determine sensitivity, a prospective study is required with an adequate sample of infants with HI. This is difficult given the low prevalence of permanent HI in a neonatal population (Hyde, 2016). The few studies that have performed such research had sample size limitations (Norton, Gorga, Widen, Folsom, et al., 2000; Wessex Universal Neonatal Hearing Screening Trial Group, 1998). A retrospective analysis of screening results among children diagnosed with HI is one method to estimate the sensitivity (Watkin & Baldwin, 2012), notwithstanding the fact that

congenital and delayed-onset HI cannot be definitively separated retrospectively (Weichbold et al., 2006).

Because of the complexity of the childhood hearing screening programme, it is necessary to define sensitivity and specificity as they relate to hearing screening. Davis et al. (2001) defined three levels of sensitivity in a hearing screening programme. The test sensitivity (or specificity) is based on the result of a single screening test. The screen sensitivity (or specificity) is based on the screening protocol, which typically includes multiple screening steps. Finally, the programme sensitivity is based on the results of the entire screening programme. Programme sensitivity incorporates factors such as coverage rate (i.e., the percentage of infants screened in a population) and the loss to follow-up (LTFU) rate (i.e., infants who are referred from screening but do not attend the follow-up appointment). Programme sensitivity therefore calculates the total number of children in a population who are missed by screening for various reasons.

2.4 NEWBORN HEARING SCREENING

It is because of NHS that the recommended 1-3-6 month benchmarks for detection, diagnosis, and intervention can be met (Yoshinaga-Itano, 2004). The following sections describe the details of an NHS programme, including the test methods, protocols, and additional programme factors.

2.4.1 Test methods

The first evidence of NHS was in the 1940s when tests were performed with noise makers in the nursery to assess the reaction of newborns (Ewing & Ewing, 1944). This practice became more standardized in the 1960s (Downs & Hemenway, 1969). The 1970s brought objective and automated methods, starting with the Crib-o-gram which measured motor activity in response to a stimulus (Simmons & Russ, 1974). It was also discovered that auditory brainstem response (ABR) was a reliable measure for estimating hearing status in infants (Hecox & Galambos, 1974). Soon after, the clinical feasibility of using OAEs in infants was revealed (Bray & Kemp, 1987; Johnsen et al., 1983). Today, ABR and OAEs are the standard technology used for NHS. These technologies are described in more detail in the following sections.

2.4.1.1 Otoacoustic emissions

Otoacoustic emissions are sounds generated from within the cochlea, produced as a by-product from the active cochlear mechanism. The emissions, once produced, travel backwards from the cochlea through the middle ear to the outer ear canal. OAEs are recorded via a probe microphone situated in the outer ear canal. They can be elicited, in response to a stimulus delivered to the ear by means of a probe microphone. Signal averaging is performed, which assumes a repeatable response against random noise. Several stimuli are presented, and responses are averaged, timed to the onset of each stimulus (Prieve & Fitzgerald, 2014).

Depending on the level of background noise or OAE amplitude, more stimuli (and corresponding responses) may be required to detect an OAE, if present (Hall et al., 1994). The resulting noise floor must be sufficiently low to observe the OAE (Prieve & Fitzgerald, 2014).

OAEs are simple to perform on newborns. The signal-to-noise ratio is not significantly affected by test location though it is affected by an active infant or background noise from an incubator (Gorga et al., 2000; Norton, Gorga, Widen, Vohr, et al., 2000). Figure 3 depicts a newborn undergoing an OAE test. The two types of OAEs used in NHS are transient-evoked OAEs (TEOAEs) and distortion-product OAEs (DPOAEs) (Hyde, 2016).

Figure 3. A newborn undergoing an OAE test. Photo by Allison Mackey.

With TEOAEs, a click stimulus is delivered to the ear. The emission produced by the cochlea is broadband. It is recorded by the probe microphone and can then be deconstructed into frequency bands. TEOAEs are highly repeatable, though their clinical application varies across frequency bands (Prieve et al., 1993). With DPOAEs, two tones (f1 and f2) are delivered via the probe speaker. Non-linear intermodulation occurs between the tones, resulting in a series of distortions (Kemp, 2002). The energy produced by these distortions is propagated back to the ear canal. Similar to TEOAE, a frequency analysis is performed, and the intensity of the distortion at 2f1-f2 is selected for analysis (Kemp, 2002). DPOAEs are plotted on a DP-gram by frequency for f2.

2.4.1.2 Auditory brainstem response

The ABR is an electrophysiological response to an acoustic stimulus generated by the auditory nerve and brainstem. ABR is a type of auditory evoked potential. Electrodes are placed on the face and head of the infant with an earphone in the ear (Jewett et al., 1970). An example configuration is shown in Figure 4. The potential difference is then recorded to

create an electroencephalogram (EEG). Similar to OAEs, the ABR is small in amplitude relative to the recorded noise. It is only observed after signal averaging. Because of the reliable nature of the ABR and assumed randomness of noise, the ABR (if present) can be identified after averaging numerous stimulus-timed EEG sweeps (Jewett et al., 1970). There is a direct correlation between the ABR threshold and behavioural audiometric thresholds (Hecox & Galambos, 1974; Sininger et al., 1997). Therefore, a minimum stimulus intensity can be selected for screening to achieve a targeted degree of HI.

Figure 4. A newborn undergoing an aABR test. Photo courtesy of Andrea Bussé.

2.4.2 Protocol

Passing criteria for both OAE and ABR screening typically use signal to noise ratios (SNR). For OAE, the SNR (in dB) is calculated for each frequency band. A minimum SNR is typically used for passing criteria together with a minimum number of frequency bands (Hussain et al., 1998; Norton, Gorga, Widen, Folsom, et al., 2000). Traditional ABR is analyzed using clinical observation of the ABR waveform. For screening, an automated approach was developed to remove the subjectivity of identifying an ABR waveform. The aABR relies fully on the internal SNR algorithm such as fine-structure processing for determining the presence or absence of a response (Sininger, 1993).

The technology and the test sequence vary across NHS programmes (Sloot et al., 2015). Figure 5 shows one example of an NHS protocol for well infants and infants in the NICU. Because of the higher prevalence of ANSD among infants in the NICU compared to well babies (Berg et al., 2011; Berg et al., 2005), the Joint Committee on Infant Hearing has recommended the use of aABR as the primary screening technology for all infants admitted to the NICU (Joint Committee on Infant Hearing, 2007, 2019). An additional screen with OAE technology for this subgroup has been recommended by some (Berg et al., 2005), as

OAE may detect cases of mild HI that were missed by aABR (Johnson et al., 2005; Levit et al., 2015). However, the Joint Committee on Infant Hearing has warned of the potential implications for higher referral rates if both technologies are used for this subgroup (Joint Committee on Infant Hearing, 2019).

NHS protocols are often multi-step sequences. As shown in the example in Figure 5 for a well-baby protocol, infants who are screened and fail step 1 are referred to step 2. If they fail again, they are referred to step 3 before continuing to diagnostic assessment. A multi-step sequence is recommended by some to reduce false positive rates (Clemens & Davis, 2001) caused by residual fluid in the middle ear or debris in the outer ear canal after birth (Kemp & Ryan, 1991).

It is important to note that the definition of a "step" is not consistent across studies in the literature, nor in the studies of this thesis. In studies I to III, a step is defined by the occurrence of screening which ends in a documented result (pass or fail). If screening is repeated immediately before a result is documented, these multiple screens are included in the step. Once a pass or fail result is documented, the step is complete. The next step may occur a few hours up to a few weeks after initial screening. In contrast, a step in study IV ends when the family leaves the location where the screening takes place. For example, all inpatient screens are counted as step 1, whereby multiple screens may be performed within hours of each other before discharge from the maternity ward. This may even include screening with different test methods.

With the exception of aABR for NICU, there is no best-practice for setting up an NHS protocol (Joint Committee on Infant Hearing, 2019). Figure 5 offers just one example; however the technology used (OAE versus aABR) and the sequence of testing varies across programmes (Sloot et al., 2015). Because of this variation in protocol design, it can be difficult for EHDI leaders to recognize the optimal sequence or technology for their local context.

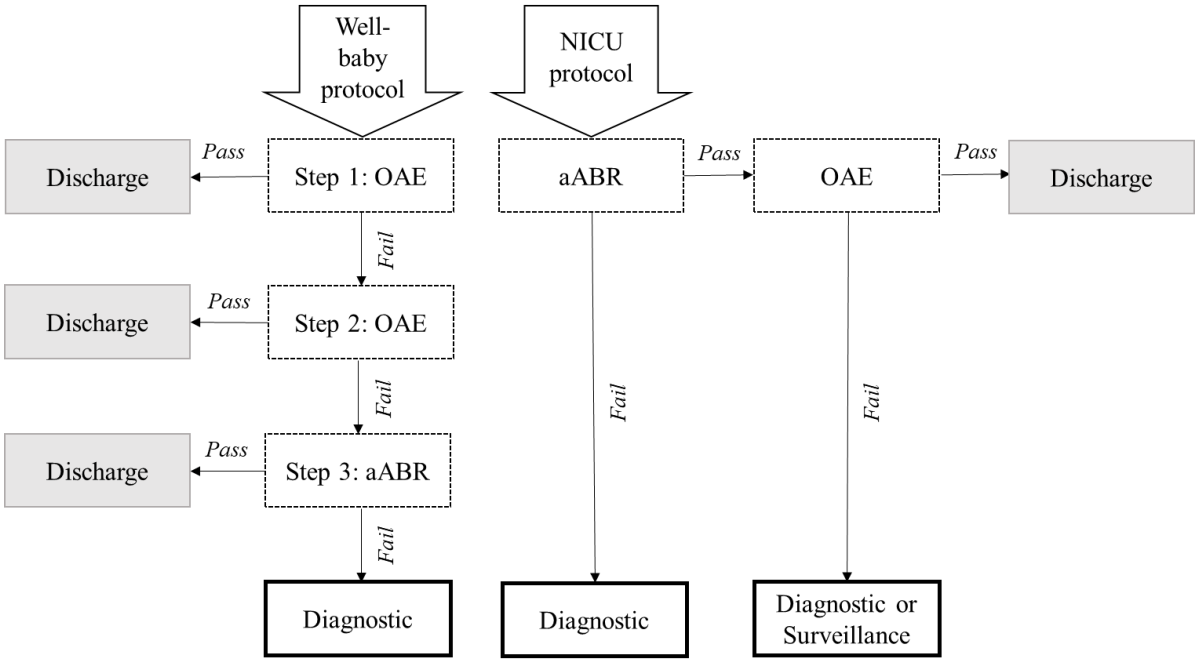

Figure 5. An example multi-step protocol for newborn hearing screening

2.4.3 Programme factors

There are other factors involved in setting up an NHS programme besides its protocol. A selection of these factors is introduced here. First, the screening personnel are the individuals performing the screening, who may be nurses, audiologists, or technicians (Cunningham et al., 2018; Stewart et al., 2000; Thomson & Yoshinaga-Itano, 2018). Second, the location of screening may be in the maternity hospital, in a well-baby clinic, or in the family's home (Fan et al., 2010; Uilenburg et al., 2009). The procedure for scheduling rescreening and the location of rescreening are also important considerations (Cunningham et al., 2018; Thomson & Yoshinaga-Itano, 2018; Uilenburg et al., 2009). Third, a minimum age of the infant at screening is often considered within an NHS programme because of the risk of false positives due to middle ear fluid or ear canal debris (Berninger & Westling, 2011; Vohr et al., 1993). Finally, various organizational factors may include whether screening for NICU and well-babies are managed under the same or separate NHS programmes (Barker et al., 2013), and whether the NHS programme is governed under local, regional, or national authority (Park et al., 2020).

2.5 CHILDHOOD HEARING SCREENING AFTER THE NEWBORN PERIOD

Childhood hearing screening after the newborn period refers to any hearing screening offered from 6 months to 7 years of age. Because the distraction or BOEL test has been phased out since NHS (Wood et al., 1997), the focus of this section is early childhood hearing screening from pre-school to school-entry age (3 to 7 years old).

2.5.1 Test methods

Methods for preschool and school-entry hearing screening include the whisper test, pure-tone (play) audiometry, and tests using alternative stimuli such as speech-in-noise or other sounds.

The whisper test is easy to perform and requires no equipment. An examiner stands behind the child and whispers either a short sequence of letters and numbers or a two-syllable word, which the child either repeats or points to the matching picture (Pirozzo et al., 2003). There are issues with the whisper test in terms of its reliability given its dependance on the level of whispered voice from the examiner (Pirozzo et al., 2003). Furthermore, it may miss a large proportion of children with mild HI (Prescott et al., 1999). Notwithstanding these limitations, it is a reasonably effective screening test, and may be useful for universally screening children in settings without access to (or resources for) audiometric screening equipment (Pirozzo et al., 2003; Prescott et al., 1999).

The most common screening test for this age group is pure-tone audiometry with play (Yong, Panth, et al., 2020). The use of pure-tone audiometry for screening dates back to the 1950s (Ewing, 1955). Pure tones are delivered via headphones using a screening audiometer. During a short conditioning session, a suprathreshold tone is presented and children are instructed to perform a task in response to hearing the tone (e.g., put a block on a peg). Once the task is understood, a tone at the screening intensity is presented. This is repeated across a frequency range (Ewing, 1955). Pure-tone audiometry screening is highly reliable and sensitive once screeners are adequately trained and the children understand the task (Lowell et al., 1956); however, it requires time and resources for equipment that needs annual upkeep. Both the whisper test and pure-tone audiometry require screening training, sustained attention and active participation from the child, and a quiet test environment.

Other tests have been created in attempt to address the limitations of the whisper test and pure-tone audiometry screen. For example, the Digit Triple Test is reliable and sensitive for detecting HI in older children and adults (De Sousa et al., 2020; Smits et al., 2013). A series of digits are presented in noise at various SNR intervals (Smits et al., 2013). The test is not susceptible to regular background noise in the room, does not require calibration, and can be performed via smartphone or tablet (Potgieter et al., 2016); however, it is not currently feasible for preschool or school-entry ages due to the task difficulty. An adapted version using ecological sound stimuli and associated pictures is under validation (Denys et al., 2019).

2.5.2 Protocol for pure-tone audiometry screening

In pure-tone audiometry screening, passing criteria are defined by the response of the child to a certain intensity of stimuli at specified frequencies. If passing criteria are not met, children may be referred to an audiology clinic for a diagnostic assessment, or first to a rescreening appointment prior to further referral (Yong, Panth, et al., 2020).

The protocol for performing pure-tone audiometry screening varies across programmes (Yong, Panth, et al., 2020). Even within certain countries, the protocol for screening varied across regions, including the frequencies tested, the minimum stimulus intensity tested, and whether a rescreening appointment was booked prior to referral for diagnostic assessment (Cadena et al., 2021; Sekhar et al., 2013; Stenfeldt, 2018). For example, in Sweden prior to nationwide standardization, Stenfeldt (2018) found out of the 12 regions that reported their passing criteria, there were nine unique combinations of passing criteria.

2.5.3 Programme factors

Together with the test and protocol, selected programme factors make up a complete hearing screening programme for preschool or school-entry aged children. First, the screening professionals may be students, nurses, audiologists, or physicians (Yong, Panth, et al., 2020). Second, the screening location could be a school (Yong, Panth, et al., 2020), well-child clinic (Stenfeldt, 2018), a pediatricians office or audiology clinic. A third consideration is the age of the child when screening is performed. Finally, the organization of screening may include whether screening is governed by the education system or public health system.

2.6 EVALUATING HEALTH PROGRAMMES

There are various conceptual frameworks that have been developed for guiding programme evaluation. Realist evaluation is a framework for complex evaluations of existing health programmes (Marchal et al., 2012). Within a realist evaluation, a health programme can be described in terms of what works (and how), for whom, and in what conditions (Pawson & Tilley, 1997). It centers around three concepts: the mechanisms, the context, and its outcomes.

The *mechanisms* are the processes within a programme (Pawson & Tilley, 1997). This could be the components, features or actions within a programme and the reasons behind their use (Dalkin et al., 2015). For example, this may be the screening method used or the procedure for scheduling a follow-up appointment. Deciding to use certain features or carry out certain actions could be influenced by the context. *Context* describes the setting, organization and conditions in which the programme is implemented (Pawson & Tilley, 1997). For childhood hearing screening, this may include staffing in maternity wards or well-baby clinics, education and experience of screening personnel, audiology knowledge of management, existing policy and practice, funding and resources, government health priority setting, or geography. Context may also include attitudes toward screening, beliefs about people with HI, or opinions for health care priority setting (Olusanya, 2015).

Together, the mechanisms and context produce outcomes. *Outcomes* are mostly discussed in this thesis in terms of quality indicators. *Quality indicators* offer a quantifiable overview of screening programme performance. The quality indicators discussed in studies I-IV are described in Table 3. They are often used in childhood hearing screening programmes and are

derived from established guidelines (Joint Committee on Infant Hearing, 2007, 2019; Wood et al., 2015). In addition to these quality indicators, study V will examine false negatives from NHS.

Table 3. Quality indicators used to evaluate hearing screening programmes in studies I-IV.

Quality indicator	Definition	Hearing Screening	Study
Coverage rate	Percentage of children screened out of the target group (complete / step 1)	- Newborn - Preschool/school-entry	II, III
Referral rate	Percentage of children referred from screening out of those who were screened (all steps / step 1)	- Newborn - Preschool/school-entry	II, III, IV
Follow-up (or lost to follow-up) rate	Percentage of children who follow-up (or do not follow-up) out of those who referred from screening (between steps or from screening to diagnostic assessment)	- Newborn - Preschool/school-entry	II, III, IV
Detection rate	Percentage of children screened who are diagnosed with permanent HI out of those screened	- Newborn - Preschool/school-entry	I, III

2.7 RATIONALE FOR THESIS

NHS has become standard practice in many countries around the world; however, there is variation in how screening is delivered (Sloot et al., 2015). There is also variation in the outcomes from screening (Kanji et al., 2018; Ravi et al., 2016). This relationship has not been comprehensively evaluated. A comparison of outcomes from quality indicators across existing NHS programmes and the previously published literature will discern the relationship between protocol and programme factors and its outcomes, while considering how the contextual factors may influence this relationship. Furthermore, since the extensive implementation of newborn hearing screening around the world, perspectives on the applicability of preschool and school-entry hearing screening are unclear. Findings can provide EHDI decision makers the knowledge they need to modify or implement childhood hearing screening based on the circumstances in their country or region.

3 RESEARCH AIMS

The aim of this thesis was to explore the mechanisms and outcomes of childhood hearing screening, considering the context in which screening programmes are implemented.

The specific objectives were:

1. To map the provision of childhood hearing screening programmes across countries or regions in Europe (Study I and III)
2. To investigate the availability and validity of data across childhood hearing screening programmes from key quality indicators used to monitor screening outcomes (Study II and III)
3. To evaluate the variation in protocols and programme factors in childhood hearing screening, considering the contextual factors across countries (Study I, II and III)
4. To evaluate the variation in referral rate and follow-up rate, with respect to protocol and programme factors (Study II and IV)
5. To evaluate the false negatives from a screening programme and assess the effectiveness of TEOAE passing criteria (Study V)

4 METHODS

This thesis comprised five studies on childhood hearing screening which used various designs, data sources and analyses. An overview of the methods for each study is presented in Table 4.

Table 4. An overview of methods for each study comprising this thesis

	Study I	**Study II**	**Study III**	**Study IV**	**Study V**
Design	Ecological	Ecological	Ecological	Systematic Review	Case-controlled
Primary data source	EUSCREEN questionnaire	EUSCREEN questionnaire	EUSCREEN questionnaire	Literature	Registries
Screening age	Newborn	Newborn	Preschool /school-entry	Newborn	Newborn
Quality indicator(s)	N/A	-Coverage rate -Referral rates -Follow-up rate -Detection rate	-Coverage rate -Referral rate -Follow-up rate -Detection rate	-Referral rate from step 1 -Lost to follow-up rate	-False negative rate
Processes, protocol or programme factors(s)	- Risk factor vs. universal - Age of infant - Test method - Steps - Benchmark ages - Passing criteria - Screening profession - Screening location	-Age of infant -Test method -Steps	-Age of child -Test method -Passing criteria -Screening profession -Screening location	Various (*explored in the study*)	-Passing criteria
Contextual factor(s)	-GDP per capita -Health expenditure -Human development index - World Bank income status -Other	- World Bank income status - Newborn care indicators/trends	-School and preschool attendance -Well-child health care routines and coverage		
Analyses	Descriptive synthesis & cluster analysis	Descriptive statistics & risk ratios	Descriptive statistics and synthesis	Descriptive synthesis & risk ratios	Multivariate conditional logistic regression

Studies I-V of this thesis made up an independent line of research in the EUSCREEN project, a large-scale collaborative Horizon2020 project with partners across the European Union that ran from 2017 to 2021. The purpose of the EUSCREEN project was to create a cost-effectiveness model and toolkit for childhood hearing and vision screening programmes. The model is publicly available and allows decision-makers to input various protocols and programme factors to design a screening programme, given the unique context in their country or region. The data collected in studies I-IV were used to inform the cost-effectiveness model and the development of the toolkit.

4.1 DATA SOURCES

4.1.1 Study I, II and III: EUSCREEN questionnaire

The EUSCREEN questionnaire was formulated by a group of hearing and vision screening experts (https://www.euscreen.org/questionnaire/). At the time of data collection, the questionnaire was published online and accessible via a unique username and password. The hearing section was made up of 191 items on hearing screening, follow-up, and intervention. Items included questions on screening tests and protocols, location and personnel for screening, background information on the organization of the screening programme, and performance indicators for screening. Most of the items were followed by a sub-item regarding the quality of the answer provided (i.e., if the answer was based on an estimation or actual data).

If actual data were used, source material and supporting documentation were requested if available. In addition, an informal search of the literature was performed using Google and Google Scholar to gather available information about hearing screening for each country or region. This included reports, documentation, presentations, media reports, and other gray literature that may support or contradict the information supplied in the questionnaires.

Data were also aggregated from public databases to reflect on the choice of protocol and programme factors used by various countries or regions. Examples of these indicators include health expenditure per capita, preschool attendance rates, and percentage of births that take place in hospital.

4.1.2 Study IV: Peer-reviewed literature for systematic review

Study IV was a systematic literature review. Two outcomes were investigated, referral rate and lost to follow-up rate from screening step 1. Peer-reviewed articles were searched across five databases (Medline Ovid, Embase, Cochrane Library, Web of Science Core Collection and Cinahl). The studies included in study IV analyzed the relationships between NHS protocol or programme factors and one or both outcomes.

4.1.3 Study V: Registries and audiogram journal

Study V linked data from the following registries and audiogram journal database:

- AudioHab, operated by Child Hearing Habilitation Center at Karolinska University Hospital Children, includes children living in Region Stockholm who are enrolled for intervention for HI. Data includes the enrolment date for intervention and degree of HI.
- Audioscreen, operated by the Center for Hearing and Balance at Karolinska University Hospital, includes all children who were born in Region Stockholm. Data includes the NHS results and TEOAE datafile.

- AuditBase (Auditdata A/S, Copenhagen Denmark), used by the Center for Hearing and Balance at Karolinska University Hospital to manage and store audiogram data.

4.2 DATA COLLECTION AND EVALUATION

4.2.1 Study I, II, and III

An ecological study was performed for studies I, III and III using responses collected from the EUSCREEN questionnaire on childhood hearing screening. The following sections briefly describe the respondents, so-called Country Representatives, and the collection and validation of their responses.

4.2.1.1 Country Representatives

An international collaboration of experts in hearing screening, vision screening and preventative child health care made up the Country-Committees Joint-Partnership of EUSCREEN Foundation. These experts, so called Country Representatives, were recruited via national and international professional organizations, networks, and peer-reviewed journal publications. After a vetting process performed by EUSCREEN project administration, Country Representatives were selected to represent their screening programme. Country Representatives were actively recruited for countries in Europe; however, no restrictions were applied on participation from other countries. In countries where only regional data were available, Country Representatives were asked to provide responses that represented their regional hearing screening programme, instead of the entire country. Collectively, the Country Representatives for hearing screening were a co-author in studies I, II and III.

4.2.1.2 Verification and validation of responses

Each response went through a multiple-step verification procedure, as follows: (1) answers were checked for completeness; (2) answers were checked across similar questions in the questionnaire; (3) answers were checked across any supporting material; (4) a preliminary draft of a Country Report was drafted, describing the details and outcomes from the hearing screening programme in a narrative form; (4) any inconsistent responses or missing information was addressed in a list of clarification questions sent to the Country Representative; (5) once all responses were received, the Country Report was updated and sent to the Country Representative for verification; (6) any comments received from the Country Representative were addressed and finalized in the Country Report. Data for analysis were then extracted from each Country Report.

For the quality indicators, data that were rough estimations based not on real data were considered ‘not valid.’ Responses that were considered valid were: (1) based on data less than 5 years old; (2) included a representative sample from the country or region; (3) were consistent with the source material and supporting documentation; and (4) contained a sufficiently large population of infants from which the data were based, i.e., a minimum of 1000 for referral rate and 5000 for detection rate.

4.2.2 Study IV

Study IV was a systematic literature review which followed the PRISMA framework for systematic reviews and meta-analyses (Page et al., 2021).

4.2.2.1 Literature sorting and selection

For the search, MESH terms and keywords were used that centered around three concepts: hearing, screening, and newborn. Three independent reviewers sorted the records by title and abstract. The inclusion criterion was that the title or abstract included reference to newborn hearing screening. The exclusion criteria were: (1) non-English report, (2) not peer reviewed, (3) screening not performed on newborns, (4) screening not for HI, (5) screening methods performed on children already diagnosed with a HI.

Full-text reports were sorted by two independent reviewers. The inclusion criteria were that the report evaluated one or more programme or protocol factor, and that one or more outcomes were reported. The exclusion criteria in the full-text review were: (1) no original data were reported, (2) screening was not performed with OAE or aABR, (3) the number of infants screened was not reported, (4) only infants with a particular condition were screened, (5) infants were older than 6 months, (6) results were described by ear and not by infant, (7) less than 100 infants were screened, (8) the report was a case study (i.e., not comparative between NHS protocol or programme factors). After sorting, discrepancies were discussed between reviewers until a consensus was reached.

4.2.2.2 Quality evaluation

Adapted from the Newcastle-Ottawa Scale (Wells et al., 2000) and QUADAS-2 (Whiting et al., 2011) a list of criteria was used to evaluate the included reports. Four criteria that assessed the risk of bias were essential for inclusion in further syntheses: (1) sampling bias was not introduced, (2) co-intervention bias was not introduced, (3) outcomes were reliable, and (4) outcomes were valid. Reports without any risk of bias were further evaluated across three domains: sample, screening and outcome (Table 5).

Table 5. Quality criteria for reports evaluated in the systematic review for study IV

Domain	Criteria
Sample	The community was described from which all infants were drawn
	The sample size was ≥1000 for each group described
	The coverage rate was reported and was ≥95%
	The infants were described with respect to risk factors, NICU admission, well babies, all babies, or other
Screening	The screening protocol included a description of at least four of the five factors: (1) infant age at screening, (2) test method, (3) number of screens included in step 1, (4) test device used, (5) passing criteria
Outcome	The method for collecting data was described
	The criteria for determining loss to follow-up was described

4.2.3 Study V

Study V was a nested case-controlled study to evaluate TEOAE results among children with and without HI who passed NHS. The following sections describe participant selection and data extraction.

4.2.3.1 Participants

Children with a later-detected HI who passed NHS were identified by linking the AudioHab and Audioscreen registries. Inclusion criteria were children born from 2006-01-01 onward, who were enrolled in AudioHab by age 8 years and before December 31, 2020, and who had passed NHS. Ear-specific exclusions were made if (1) TEOAE screening data were missing from Audioscreen, (2) TEOAE data indicated that screening should have failed even though database entry indicated pass, (3) hearing threshold data from AuditBase showed normal hearing thresholds, (4) HI progressed significantly, according to the definitions of progression from Barreira-Nielsen et al (2016) and Watkin and Baldwin (2012), and (5) the HI was conductive according to the definition in Davis and Davis (2016). Randomly selected controls were extracted from the Audioscreen registry, matched to each case on date of birth, sex, and ear (left or right), at ratio of three controls per case.

4.2.3.2 Screening data

TEOAE database files were downloaded from Audioscreen, and data were extracted to ASCII files with Otodynamics ILOv6. These data included: test duration, stimulus stability, stimulus intensity, response reproducibility, total response amplitude, total noise amplitude, and response and noise amplitudes for each of the frequency bands centered around 1, 1.5, 2, 3 and 4 kHz. SNR was calculated from the response and noise amplitudes.

4.2.3.3 Hearing impairment data

Hearing thresholds were extracted from AuditBase. The first two audiograms (within a maximum of 1 year) were used to cross-check initial thresholds. If initial thresholds were marked as unreliable, the later thresholds were used. Bone- and air-conduction thresholds were extracted and used to exclude conductive HI. Four frequency pure-tone averages (500 to 4000 Hz) were calculated from the initial audiogram to identify the initial degree of HI. An audiogram 3 years later was extracted to identify progression. If not available, the audiogram closest to this 3-year time point was used.

4.3 ANALYSES

4.3.1 Study I

A descriptive synthesis was performed on the various aspects of NHS provision across the countries or regions surveyed. To evaluate the choice of NHS protocol components, five variables were selected for analysis: First, the programme type described whether there was (1) no NHS, (2) targeted NHS for high-risk infants, (3) a single-protocol design (universal NHS using one protocol for all infants), or (4) a dual-protocol design (universal NHS using

separate protocols for high- and low-risk infants). Second, the test method for low-risk infants could be OAE-only or include aABR. Third, the test method for high-risk infants could be OAE only or aABR for step 1. The fourth and fifth variables were the number of steps for the low-risk and high-risk protocols. A hierarchical agglomerative cluster analysis was performed in SPSS v. 26.0 (SPSS Inc., Chicago, IL). Using this method, clusters are formed demonstrating groups of countries or regions which were most closely aligned to their NHS protocol.

4.3.2 Study II

The analysis of study II focused on the outcomes from NHS. Quality indicators were coverage rate, referral rates, follow-up rates, and detection rate. Valid responses for the quality indicators were transformed into estimated numbers based on the birth rate in the country or region represented. Pooled rates were calculated and compared across protocol and programme factors and risk ratios were calculated.

4.3.3 Study III

The analysis of study III focused on the process, protocols, programme factors and outcomes from childhood hearing screening after the newborn period. The analysis performed in study III was a descriptive synthesis.

4.3.4 Study IV

Reports were organized by protocol or programme factors and the changes in referral or loss to follow-up rates were quantified. Because of the heterogeneity between studies, it was determined that a meta-analysis was not appropriate; however, risk ratios were performed across individual studies where sufficient data were supplied. This illustrated the increased or decreased risk for referral or risk for loss to follow-up, based on the protocol or programme factor studied, as well as the variation of these effect across reports.

4.3.5 Study V

A multiple conditional logistic regression was performed using R (Mazerolle, 2020; R Core Team, 2021; Therneau, 2022). This analysis is used to assess the variables that significantly predict the outcome of HI when several variables are simultaneously included into the model. The variables tested in the regression were the absolute response levels at each frequency band (dB sound pressure level) and whether SNR was ≥6 dB or <6 dB. Because SNR and response amplitude were interdependent, SNR was converted to a dichotomous variable so that both the SNR and response amplitude could be entered into the same analysis. There was a correlation between 1.5 and 2 kHz bands and between 3 and 4 kHz bands, which broke the assumption of collinearity for a logistic regression. Therefore, these four variables were collapsed into two: mid-frequency bands (1.5-2 kHz) and high-frequency bands (3-4 kHz). The logistic regression was performed using a backwards stepwise elimination method and the best model fit was selected according to the Akaike information criterion. Other analyses

performed in study V included t-tests and chi-squares to illustrate differences between groups of ears with and without HI.

4.4 ETHICAL CONSIDERATIONS

Of the studies in this thesis, only study V required ethical approval. Studies I, III, III and IV did not use personal data. Nevertheless, there were still ethical issues to consider for these studies. The ethical considerations of bias and avoiding harm will be discussed with regards to Studies I-IV. The ethical considerations of using personal data will be discussed for study V.

The EUSCREEN questionnaire aggregated performance indicators for 47 national and regional screening programmes. Country Representatives were typically leaders in pediatric audiology, experts or managers of the local hearing screening programmes. Outcomes from screening programmes were published as part of study II and compared to other programmes. The first ethical dilemma was with regards to reporting bias. Naturally, the EUSCREEN project may have positioned Country Representatives in a potentially vulnerable situation. In the EUSCREEN studies, performance data representing each Representative's local screening programme was published in a comparative evaluation. Therefore, one may speculate on the potential for falsified or skewed reporting on the part of the Country Representative in order avoid negative impact from the comparison. This issue is mitigated in several ways. First, the Country Representatives were valued collaborators in the project, and the study was carried out with full reciprocal transparency. A mid-project meeting was held in March 2019 in which preliminary results were discussed with all Country Representatives in attendance. Meetings were also held at other international conferences and virtually throughout the project's timeline (2017-2021). Country Representatives were collectively a co-author in Studies I, III and III, and therefore were held accountable for their contributions. Second, the systematic process by which all responses were cross-checked, verified, and clarified reduced the likelihood for reporting bias. Any responses which seemed inconsistent were questioned and clarified. Only data which met high standards for validity were reported in the study. Finally, the purpose of the EUSCREEN project was very clear. It was to develop a cost-effectiveness tool in which programmes could be designed to produce the optimal configuration given any unique context. It was therefore required that the relationship between the mechanisms, context, and outcomes be assessed. The purpose was not to compare programmes for the goal of placing judgement. In order to avoid harm in this sense, it was important to carefully consider the wording used, particularly in study II.

Reporting bias is also a consideration for study IV in peer-reviewed literature. Similar to the EUSCREEN questionnaire, the evaluation of the quality of the reports mitigated the synthesis of any study that showed bias. This was performed through eliminating reports from analyses if they did not meet all four essential criteria regarding risk of bias.

Avoiding harm is a second ethical consideration, which is relevant for Studies I-III. The concept of avoiding harm may be with respect to the interpretation and impact of the findings

from Studies I-III. For example, if certain NHS programmes appeared to be performing poorly, this could potentially have negative implications for stakeholders. Again, it was not the objective of the EUSCREEN study to compare performance measures in this sense, but only to disseminate findings that might improve childhood hearing screening in all countries.

Study V is the only study that used individual level data. Ethical approval was granted by the Swedish Ethical Review Authority (2020-07302). All registry studies are faced with ethical dilemmas regarding integrity and autonomy. The first issue is the importance of safeguarding personal integrity. In this research, personal data are used about HI in children. There is always the risk of a breach in clinical based research; however, with proper measures in place to safeguard this data, this risk was minimal. All data were pseudonymised and stored on a secure and institute-approved server for personal data usage. Second, there is the issue of autonomy. As is the case with many registry-based studies, informed consent was not requested or required. To answer important research questions in this study, data from the entire target cohort was needed. If informed consent was required, the reduced power would make the study not possible and selection bias would be introduced. The benefits of being able to carry out this study outweighed the risks.

5 RESULTS AND DISCUSSION

The main findings from this thesis are summarized in Table 6. More detailed results are presented in the subsequent sections, combined with a discussion.

Table 6. Summary of main findings from studies I through V.

Study	Main Results
I	• NHS programmes existed in 42 out of 47 countries or regions surveyed • Out of 30 high-income countries, 25 had nationwide universal NHS; out of 15 low- or middle-income countries surveyed, 3 had nationwide universal NHS • Using variables on the status and protocol for NHS, five clusters were optimally generated from the cluster analysis. They were named: (1) n*o NHS*, (2) *targeted NHS*, (3) *single-protocol for all infants with OAE only*, (4) *dual-protocol with OAE-only for low-risk infants*, (5) *dual-protocol with aABR included for low-risk infants*. Dual-protocol design means different protocols were used for low- and high-risk infants. • Countries or regions in the *No NHS* cluster had lower health expenditure per capita and human development index compared to the two clusters with a *dual-protocol design*. • Countries or regions in the cluster *single-protocol design with OAE only* had lower health expenditure per capita and human development index compared to countries in the cluster *dual-protocol design with aABR for low-risk infants.* • The most common protocols for low-risk infants were a 2-step OAE-OAE protocol and a 3-step OAE, OAE, aABR protocol. • The most common protocols for high-risk infants were aABR only or OAE+aABR. Ten countries used only OAE for high-risk infants. • Most Country Representatives did not report the passing criteria for OAE, citing that they used the default algorithm in the screening device. Passing intensity for aABR ranged from 30 to 45 dB nHL with the majority of programmes using 35 dB nHL. • The most common age of the infant at screening was 24 to 72 hours after birth. • Nurses were the most common screening professional; other professionals included midwives, audiologists, physicians, and technicians/screeners. • Most screening took place in the maternity hospital before discharge. Other screening locations were a well-child clinic/centre or the family's home.
II	• Out of 42 NHS programmes, a valid coverage rate was reported by 26 programmes, referral rate by 22 programmes, follow-up rate to diagnostic assessment by 12 programmes, and detection rate of permanent HI by 13 programmes. • Most programmes reported coverage rates of 95% or higher. There was variation in how coverage rate was calculated across programmes.

	• Referral rate for low-risk infants or all infants combined was below 4% for all programmes that reported valid figures. For high-risk infants, referral rate ranged from 4% to 10%. • Follow-up rate to diagnostic assessment ranged from 19% to 97% across 12 programmes providing valid figures. Eight of these 12 had follow-up rates under 90%. For nine programmes, the follow-up rate after step 1 ranged from 27% to 97%. • Referral rate from step 1 ranged from 1.8% to 15.3% for programmes that screened after 24 hours from birth. For two programmes that screened earlier than 24 hours, referral rates of 22% and 6% were reported. For programmes that screened after 72 hours from birth, referral rates were 4%. • Programmes that used a 1 or 2 step OAE-only protocol for low-risk infants, the pooled referral rate was 2.1%. This is compared to programmes that used a 2-step protocol with aABR (pooled referral rate of 1.7%) and programmes that used a 3 or 4 step protocol with aABR (pooled referral rate of 0.8%).
III	• Out of 47 countries or regions surveyed, 17 provided universal preschool- or school-entry hearing screening; an additional eight provided limited screening (opportunistic, project-based, or localized to certain regions) • Out of the 17 universal programmes, all used pure-tone audiometry screening. Out of the eight limited programmes, six used pure-tone audiometry screening and two used the whisper test. • All except two of 17 programmes were implemented prior to NHS. Since NHS implementation, two programmes have scaled down or cancelled screening. One country halted any new implementation until more data becomes available on its effectiveness. • The age of the children when screened ranged from age 3 to 7 years. Screening was performed in a school setting or a well-child clinic or both. Nurses were the most common professionals to perform screening; other screening professionals were speech-language therapists, audiologists, or physicians. • Three Country Representatives provided valid data on the coverage rate in their countries (97%, 99% and 45%). A valid referral rate was available for two countries (7.6% and 7.9%). Follow-up rates to diagnostic assessment were 58% and 77% across two countries. The detection rate of permanent HI was 0.012% from one country.
IV	• 6047 records were identified from the search after duplicates were removed; 1801 records met title and abstract criteria and full-text reports were screened; a total of 101 studies were included in the synthesis. • For studies on LTFU, 30% reported a coverage rate over 95%; 53% reported the method of data collection; and 35% reported the definition of LTFU. For

	studies on referral rate, 16% reported a coverage rate over 95%; and 44% reported the method of data collection. • Thirty-five studies reported on both the referral rate and LTFU rate from step 1; 58 studies reported on referral rate only; eight studies reported on LTFU rate only. • Most studies reported lower referral rates from step 1 when using aABR compared to TEOAE; however, there was a wide range of referral rates for each test method (3% to 71% for TEOAE and 1% to 23% for aABR). • Passing criteria for OAE did not significantly affect referral rate from step 1. • Rescreening immediately after a failed result or before discharge from the maternity hospital resulted in an overall lower referral rate at discharge. • Screening only when the infant is calm lowered the referral rate from step 1. • Screening with OAE on day 3 to 5 after birth resulted in the lowest referral rate, compared to before day 3 and compared to weeks to months after birth. • Overall, more experienced screeners (i.e., those that screened more infants) had lower referral rates and lower LTFU rates; however, this trend was mitigated by other elements, such as a network with larger hospitals or an audiologist on staff. One study was the exception because screeners working in larger hospitals were likely overburdened. • The location and fees for step 2 were important factors in the risk of LTFU. • High-risk infants appeared to be more at risk for LTFU, particularly if LTFU was relatevely low overall. This finding was not consistent across all studies. There was too much heterogeneity for a quantitative meta-analysis.
V	• From a total of 1244 children with HI in the cohort, 260 were registered as having passed TEOAE screening in both ears, out of which 211 had TEOAE data available. • 41 out of 211 children were incorrectly entered into the database as having passed TEOAE screening, out of which 24 had no follow-up data in the screening registry. • In the final regression model, the odds of a HI were decreased by a factor of 4 if SNR at 4 kHz was 6 dB or more. The odds off a HI decreased by 10% for each 1 dB increase in TEOAE amplitude in the mid-frequencies (1.5 and 2 kHz bands combined). • In the sample of children with HI who passed TEOAE, there is a direct correlation between pure-tone average threshold and TEOAE amplitude for the 3 and 4 kHz frequency bands, in that mild HI had lower TEOAE amplitude than more severe HI.

5.1 MAPPING THE PROVISION OF CHILDHOOD HEARING SCREENING (Study I and III)

Country Representatives from a total of 47 countries or regions completed the EUSCREEN survey. In accordance with the first objective, the provision of childhood hearing screening was mapped, primarily across Europe. This included newborn hearing screening and childhood hearing screening after the newborn period.

5.1.1 Newborn hearing screening

Results showed that 42 participating countries or regions have implemented NHS to some degree. Five out of the 30 high-income countries that participated lacked universal nationwide NHS. For the 15 low- or middle-income countries (LMICs) that participated, 12 lacked universal nationwide NHS. Though in these countries, some screening was performed. This was either targeted NHS, universal NHS in some hospitals or regions, or research-funded screening projects. The cumulative number of countries in the European Union (EU) that implemented universal NHS as of 2018 is displayed in Figure 6, in addition to the number reaching nationwide status.

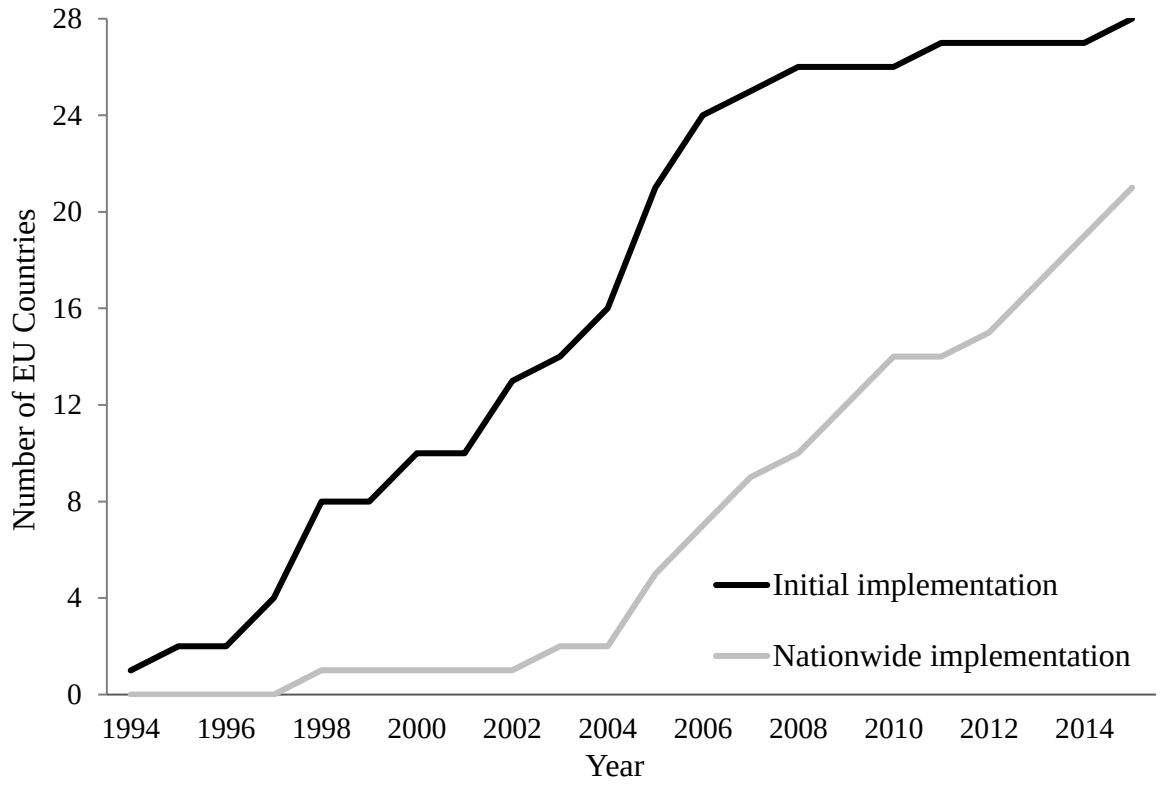

Figure 6. The trend in the year and cumulative number of countries in the European Union (as of 2018) where newborn hearing screening is implemented and achieving nationwide reach. Reprinted with permission by Taylor and Francis Group.

The cluster analysis performed in study I grouped countries or regions based on similarities and differences in NHS programme characteristics. Five clusters were generated: (1) *No NHS*, (2) *Targeted NHS*, (3) *Single-protocol for all infants with OAE only*, (4) *Dual-protocol with OAE-only for low-risk infants*, (5) *Dual-protocol with aABR included for low-risk infants*. Clusters were then compared in terms of the health economic and development status

in the countries. The *targeted NHS* cluster was excluded from this comparison because there were only two programmes in this cluster (Malta and North Macedonia). As shown in Figure 7, significant differences were found for health expenditure per capita between countries that fell into the *no NHS* cluster (n=5) and countries that fell into one of the two clusters that used a dual-protocol design (n=14 and n=17). Similarly, significant or near significant differences were observed in the human development index between the *no NHS* cluster and the two clusters that used dual-protocol designs.

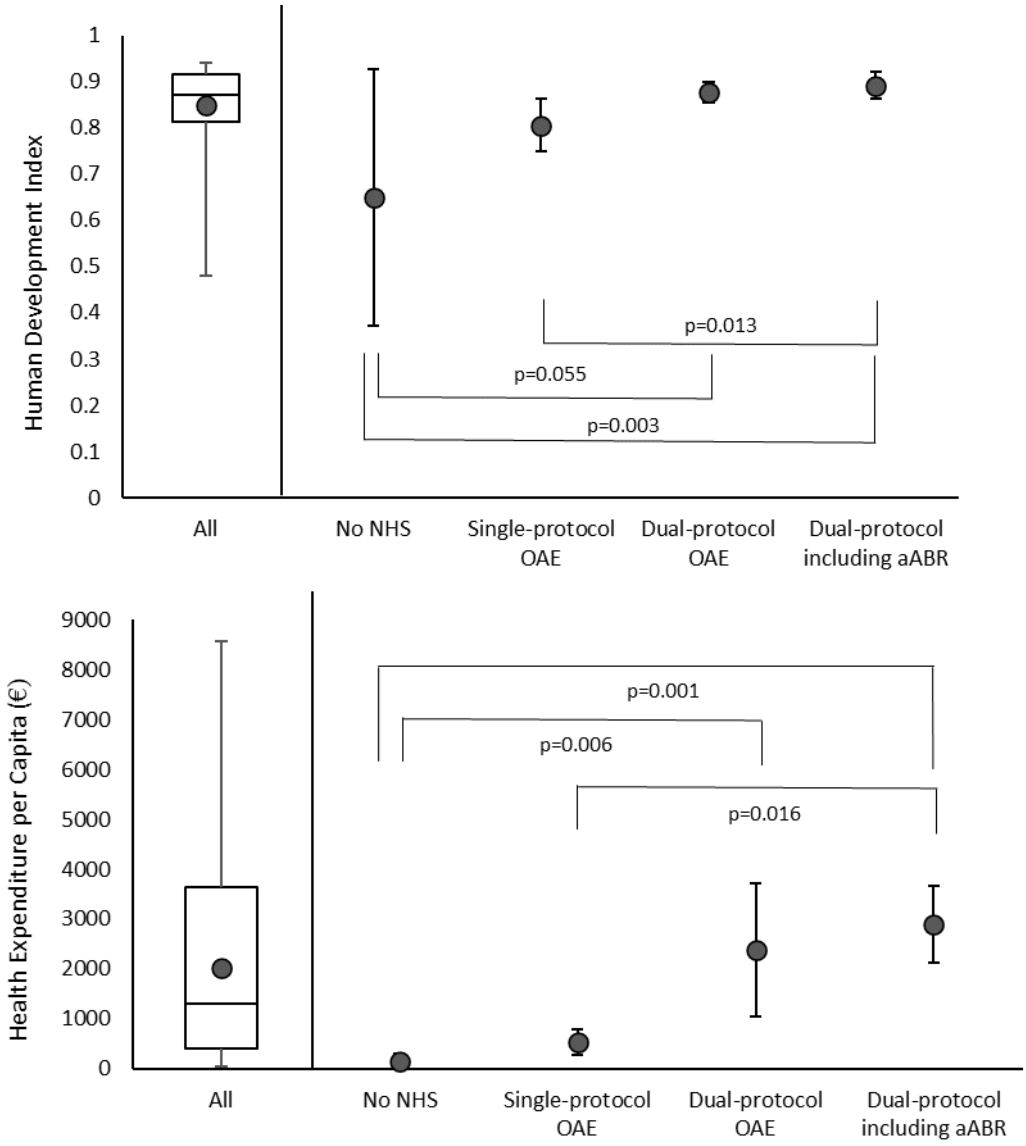

Figure 7. Health expenditure per capita and human development index for countries divided by clusters and for all participating countries combined. Filled circles represent the mean. Error bars represent the 95% confidence interval. Significant and near significant differences were based on non-parametric post hoc comparisons and adjusted with the Bonferroni correction.

In sum, study I showed that participating countries without NHS had the lowest health expenditure per capita and human development index. Olusanya (2015) also described the case that few LMICs have NHS. In LMICs, the highest percentage of childhood HI is preventative (e.g., maternal infections, birth complications); and therefore, the World Health Organization has recommended dedicating resources toward primary measures for preventing

childhood HI over measures for detection (Neumann et al., 2019; Wilson et al., 2017; World Health Organization, 2016). However, efforts toward prevention will never remove all permanent HI. Therefore, even with primary measures in place aimed to reduce preventative childhood HI, secondary measures for detection, i.e., NHS, may be most critical in LMIC where prevalence rates of childhood HI are highest (Olusanya, 2012, 2015).

In a review of NHS among LMICs, Olusanya (2015) outlines some key health care and societal challenges for provision of NHS in countries with limited resources. For instance, hearing impairment may be considered a low priority among health care workers, particularly in comparison to other conditions with a risk of infant mortality. Additionally, a lack of funding for screening, diagnostic and intervention services, a lack of infrastructure for sustainability of the referral and intervention pathway, and a societal stigma toward HI may all be barriers to overcome.

In their World Report on Hearing, the World Health Organization (2021) underscores the importance of universal NHS for detecting all infants with HI. There is general agreement that universal NHS is cost-effective in high-income settings (Sharma et al., 2019; Sharma et al., 2022). In more resource-constrained settings, results are not as clear. Studies based in Thailand and Albania reported that universal NHS was cost-effective compared to no screening (Pitathawatchai et al., 2022; Verkleij et al., 2021), while a study in China reported that targeted NHS was more cost-effective than universal NHS, suggesting a gradual roll-out to universal screening once other EHDI services become more aligned (Tobe et al., 2013). The World Report on Hearing recognizes the barriers to full-scale implementation of universal NHS in resource-constrained settings. (World Health Organization, 2021). Their recommendation is that targeted or even opportunistic screening may be a feasible first step toward reaching the goal for universal and nationwide implementation (*Hearing screening: considerations for implementation*, 2021; World Health Organization, 2021).

5.1.2 Childhood hearing screening after the newborn period (Study III)

Unlike NHS, pre-school or school-entry screening is not widely available in Europe. Out of the 47 participating countries or regions, 17 reported universal preschool or school-entry hearing screening. While some countries are newly implementing preschool or school-entry programmes, others have halted the implementation of new programmes, cancelled or scaled back existing ones.

This trend across countries exemplifies the lack of clarity in the literature regarding the cost-effectiveness of preschool or school-entry screening programmes. Some studies and expert groups have advocated for the universal preschool-age or school-entry hearing screening (Lü et al., 2011; Skarżyński & Piotrowska, 2012). However, there is currently a lack of high-quality data on the effectiveness of preschool-age or school-entry hearing screening (Fortnum et al., 2016; Yong, Liang, et al., 2020). One recent study supported the implementation of a school-entry hearing screening in Australia, reporting a 96% likelihood that the programme was cost-effective (Gumbie et al., 2022). Another study showed that school-entry screening

in one region of England was not more effective at detecting children with HI (or detecting them earlier) compared to another region without screening (Fortnum et al., 2016). The authors themselves noted limitations to the study and that more research is needed. Regardless, this study has led to disinvestment of school-entry screening programmes in England, which was also described from the survey in study III.

5.2 AVAILABILITY OF DATA FROM CHILDHOOD HEARING SCREENING PROGRAMMES (Study I, II, III and IV)

It was evident from studies I, II and III that many experts involved in regional or national hearing screening programmes were unable to provide valid information or aggregate data on the performance of their programme. First, findings from studies I, II and IV illustrated a lack of continuity in the EHDI pathway from screening to intervention. Furthermore, results from studies II and III indicated a lack of systematic data collection and quality assurance procedures for many participating programmes for both newborn and preschool or school-aged hearing screening. Finally, studies II and IV revealed inconsistencies or insufficiencies regarding the definitions and documentation of the quality indicators used for measuring the performance of programmes.

5.2.1 Lack of continuity of the EHDI pathway

In the EUSCREEN survey, questions existed regarding the organisation, policy, protocols, other programme components (e.g., screening profession and location), quality indicators, costs, prevalence, and outcomes. Study I described the difficulty Country Representatives had in reporting on all parts of their programme. Figure 8 illustrates the completion rate of questions from participating countries or regions. Most Country Representatives could report on the organization and planning of NHS, and the protocol and programme components; however, for questions relating to monitoring or evaluation or intervention of HI, fewer than 20% of the Country Representatives completed all questions.

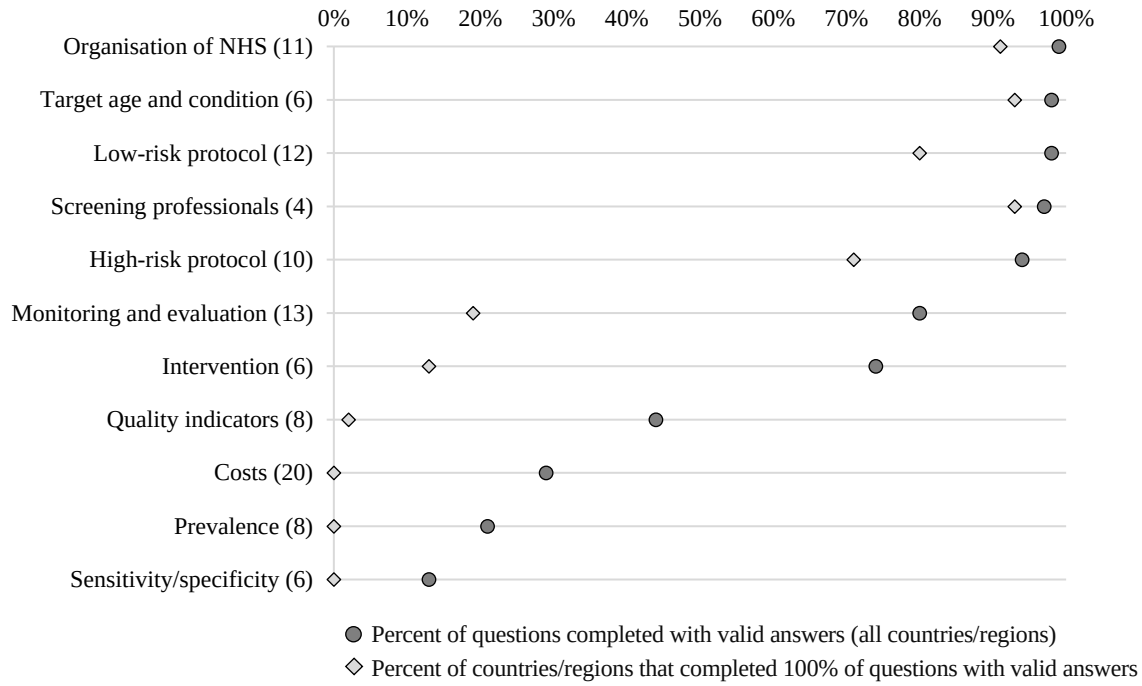

Figure 8. The completion of questions from the EUSCREEN questionnaire, divided by category. For each category, two value are represented: the percent of questions with valid answers combined across all participating countries, and the percent of countries or regions that provided valid answers to all questions in that category. The number in brackets represents the total number of questions in the respective category.

This finding was also reflected in studies II and IV. As shown in Table 7, more NHS programmes had valid data on the indicators of coverage and referral rates, while few data were available on the follow-up rates to diagnostic assessment, or the number or percentage of children detected with HI. In study IV, 58 reported only on referral rate, 35 studies reported on both referral and LTFU rate, and eight reported only on LTFU rate. These findings represent an issue among EHDI programmes described elsewhere, which is the lack of continuity in governance and data information systems between screening, diagnostics and intervention (Alam et al., 2016; Bagatto et al., 2020; Russ et al., 2010). This breakdown impacts families attempting to navigate the system, leading to LTFU between steps of the EHDI pathway, and children with potential HI detected later or not at all (Hunter et al., 2016; Russ et al., 2010). High-quality and secure data information systems are required to overcome the challenges of tracking infants from the birth hospital to follow-up and intervention (Finitzo & Grosse, 2003).

Table 7. The number (percentage) of programmes where valid data could be provided for each of the quality indicators studied.

	Study II: newborn hearing screening programmes (n= 42)	Study III: preschool/school-entry hearing screening programmes (n=17)
Coverage rate of screening	26 (62%)	3 (18%)
Referral rate from screening to diagnostic assessment	23 (55%)	2 (12%)
Follow-up rate to diagnostic assessment	12 (29%)	2 (12%)
Detection rate of permanent HI	13 (31%)	1 (6%)

5.2.2 Poor availability of aggregate data on quality indicators

While many Country Representatives could report on descriptive information about their programme, few valid data were available regarding the performance of their screening programme. As shown in Table 7 for study II, less than two-thirds of Country Representatives could provide valid data on coverage or referral rate for their programme, and less than one third could provide valid data on follow-up or detection rate. Not reported in this figure is that only 9 programmes reported on the percentage of infants referred from step 1 screening who followed up for step 2 (rescreening). For preschool or school-entry hearing screening in study III, the availability of data was even poorer than for NHS. Valid data were only available from a few programmes.

Country Representatives had two years to complete the questionnaire with ongoing communication and support from study partners, and often sought help from other local professionals. For example, in cases where NHS was administered separately from intervention services, the Country Representative delegated questions specific to HI intervention to other professionals. Still, much of these data were missing. Unfortunately, the reason that data were not available was not investigated in these studies. It is possible that data are collected internally by local hospitals or well-child units and not reported to regional or national authorities; data may be reported but not accessible for distribution; or data may not be collected at all.

Documenting and reporting outcome data are key for evaluation of a health system, be that for a systematic evaluation, a cost-effectiveness analysis, or for continuous quality improvement processes. Finitzo and Grosse (2003) recognized the immediate need for effective quality management systems during the rise of EHDI programmes. They stated that quality management systems include procedures to monitor and evaluate performance, set benchmarks for quality indicators, and then use this information to guide modifications to the programme (Finitzo & Grosse, 2003). In the United States, the value of quality management, reporting, and sharing performance data has led to a national database and an annual survey submitted by state-wide EHDI programmes (Alam et al., 2018; Gaffney et al., 2010). This comparison across programmes allows for targeted federal support based on documented need and improvement goals (Alam et al., 2018; Subbiah et al., 2018). Sharing performance

data across EHDI sites is vital for gaining external perspectives, provide feedback, and improve outcomes (Finitzo & Grosse, 2003).

5.2.3 Lack of standardization of quality indicators

Another finding from studies I to IV was the inconsistencies in how (or whether) quality indicators were defined in NHS programmes and reports. In the EUSCREEN study, the inconsistency with regards to definition, particularly for coverage rate, was noted early in the data collection process, and a glossary was distributed to Country Representatives. Part of the validation process required the Country Representatives to define their quality indicators against the glossary.

As described in study II, the definition of coverage rate was inconsistent across the 26 NHS programmes where valid coverage rate data were available. For 22 programmes, coverage rate was defined as the percentage of infants who completed step 1 screening. For the remaining four programmes, it was defined as the percentage of infants who completed all steps of the screening process. For 19 programmes coverage rate was provided for all infants and for the other three programmes coverage rates were provided separately for well babies and NICU babies. A similar variability was observed for referral rate and follow-up rate.

Despite validation efforts, in-depth scrutiny of the detailed criteria for each quality indicator was not performed in the EUSCREEN project. For example, it was unclear if families who explicitly refused screening or diagnostics were included in the coverage rate or follow-up rates. These inconsistences and inadequacies with regards to defining quality indicators also existed in reviews of the literature (Mincarone et al., 2015). The inconsistent definitions used when defining permanent HI affected the comparability of prevalence rates for newborn HI (Bussé et al., 2020). In study IV, only 35% of studies that evaluated and reported on LTFU provided a description of how this indicator was calculated in the report, which also limits the capacity to make valid comparisons across studies.

The need for standardization of definitions across programmes was lifted by the Joint Committee on Infant Hearing (2007) as it related to state-wide programmes in the U.S., where national estimates are calculated and aggregate data are compared across programmes. Differing definitions leads to under- or over-estimations in comparison to a base formula (Alam et al., 2016). A lack of consistency directly impacts the comparability of performance data and the feasibility for making quality improvements based on the comparative analysis (Mason et al., 2008). Conversely, the need for standardization may not be practicable for independent programmes that have used set definitions for their quality indicators for many years and have established internal benchmarks based on these definitions.

5.3 EVALUATION OF OUTCOMES FROM NEWBORN HEARING SCREENING (Study II, IV and V)

The following sections describe and discuss the results of outcomes of NHS collected from studies II, IV and V. The coverage rate, referral rate, follow-up rate, and false negatives are described. Results on the detection rates aggregated can be found in study II.

5.3.1 Coverage rate

Coverage rate outcomes were reported in study II. For the 22 programmes that defined coverage rate as the percentage of infants screened with step 1, five programmes reported rates less than 95%. This included Romania where a nationwide coverage of 18% was reported, as their screening programme is in the process of scaling up. For the four programmes that defined coverage rate as the percentage of infants completing all screening steps, coverage rates ranged from 93 to 99%.

5.3.2 Referral rate

Referral rate can be evaluated from screening to diagnostic assessment or for each individual step in a screening programme. In study II, the referral rates from screening to diagnostic assessment for high-risk infants ranged from 4.0% to 10.2% across seven NHS programmes. For low-risk infants, the referral rate to diagnostic assessment ranged from 0.3% to 3.5% for 14 programmes. Ten programmes reported the final referral rate for all infants, which ranged from 0.3% to 3.5%. The issue with reporting referral rate from screening to diagnostic assessment is that any loss to follow-up between steps will cause an underestimation of the false positive rate. In Tuzla Canon (Bosnia and Herzegovina), the final referral rate to diagnostic assessment was 0.4% of all infants screened; however, only 27% of infants followed up to step 2 after being referred from step 1. Consequently, the low referral rate is mostly a reflection of high LTFU between steps. As described, very few NHS programmes reported on the follow-up rate between steps 1 and 2. This issue remains a limitation in how referral rates are interpreted.

The referral rate from step 1 was also evaluated in this thesis in studies II and IV. Figure 9 illustrates referral rates reported in study II for each step 1 to 4 (where applicable) for the NHS programmes where these data were available and valid. The influence of a multi-step protocol on the final referral rate will be discussed in a later section. However, these data also demonstrate the variability in step 1 referral rates across NHS programmes. A similar trend was observed in study IV in which step 1 referral rates ranged from 0.6% to 32.2% across 41 studies on low-risk infants. The wide variability in referral rate from step 1 is further discussed in a following section of this thesis as it relates to the mechanisms and context of an NHS programme.

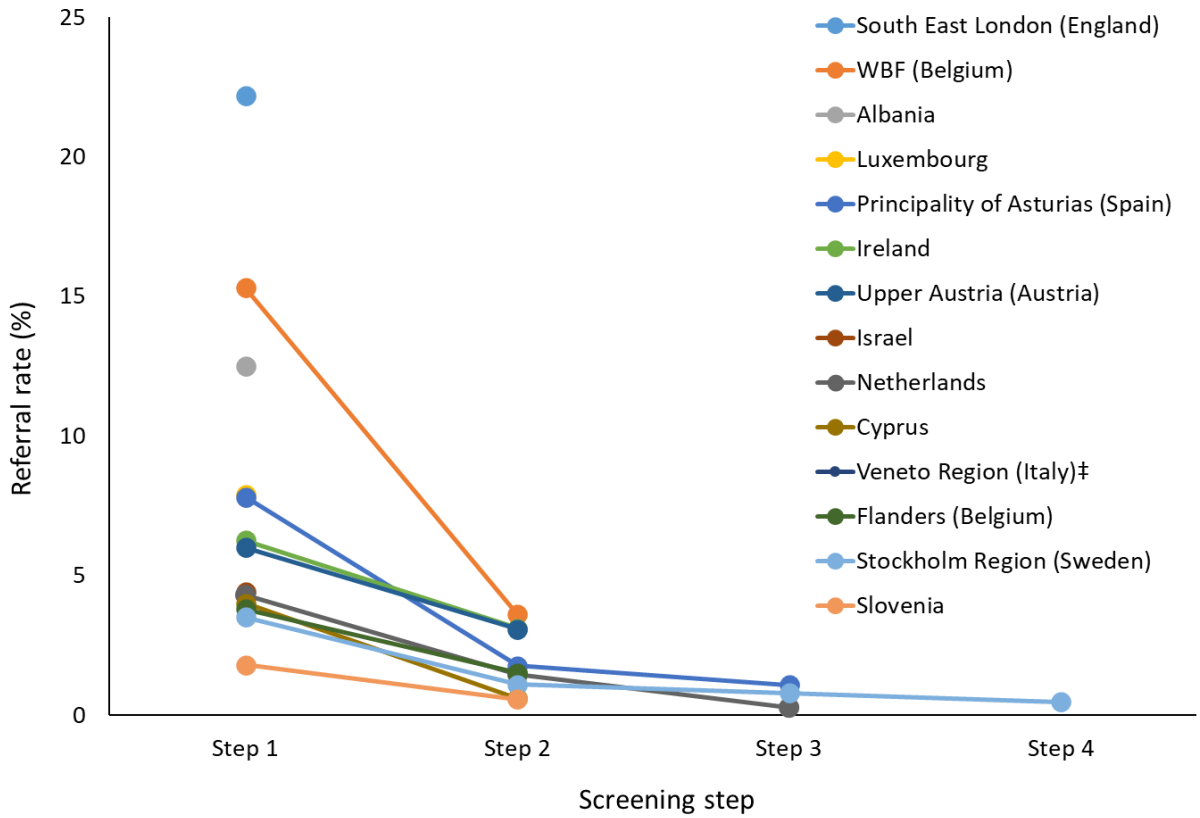

Figure 9. Referral rates for low-risk infants from steps 1 up to 4 were reported by Country Representatives, where applicable and if valid data were available. For some NHS programmes, only referral rates from step 1 were available.

5.3.3 Follow-up rate or lost to follow-up rate

The follow-up rate (or lost to follow-up rate) was analyzed in studies II and IV. Follow-up rates were quantified from step 1 to step 2 and from the final screening step to diagnostic assessment. In study II, the follow-up rate from screening to diagnostic assessment ranged widely, from 19% to 97% across 12 programmes. Only four of these 12 programmes reported follow-up rates of 90% or higher. The follow-up rate between steps 1 and 2 ranged from 27% to 97% across nine programmes. This variability was also demonstrated in study IV.

The results of this thesis showed that the problem with LTFU is ubiquitous. LTFU is clearly a major barrier for EHDI in many countries and regions. Simply put, children are referred from screening because of the suspicion of HI; if these children are not followed-up, the effectiveness of the programme is reduced (Verkleij et al., 2021). Most importantly, there is a risk that children with HI are not offered the timely intervention they need. As described in study IV, the proportion of studies analyzing programme factors that affect LTFU were lacking in comparison to studies aiming to improve referral rate. Although not included study IV, studies that do exist about LTFU often analyzed family-level risk factors, such as material smoking habits (Cheung et al., 2022; Cunningham et al., 2018; Razak et al., 2021; Zeitlin et al., 2017; Zeitlin et al., 2021).

It is particularly vital that infants with risk factors for HI are followed up. Study IV reviewed 24 reports where LTFU was measured between high- and low-risk infant groups. The reports

were inconsistent as to whether high-risk infants had higher rates of LTFU compared to low-risk infants. Two reports that modeled the influence of risk group revealed a significantly higher likelihood that high-risk infants were LTFU compared to low-risk infants (Razak et al., 2021; Vohr et al., 2002). Although these articles bring awareness about the issue and barriers to follow-up, very few studies have analyzed possible strategies for improvement (Hunter et al., 2016; Russ et al., 2010; Thomson & Yoshinaga-Itano, 2018). More focus on quality improvement strategies can help improve follow-up rates for all children referred from screening.

5.3.4 False negatives

The false negatives from screening were estimated in study V. Although a question on sensitivity and false negative rate was asked in the EUSCREEN questionnaire for study II, no Country Representative provided these data. This was expected, given the complexity in calculating sensitivity or false negatives from newborn hearing screening. In study V, false negatives from TEOAE screening were estimated based on a retrospective analysis of children with HI who passed. Out of 1244 children with HI in the cohort studied, 260 were documented as having passed TEAOE screening in both ears; 211 of whom had TEOAE data available. After cross-checking the TEOAE data to the passing criteria, it was observed that 11% of children were misidentified as having passed the screening. Furthermore, two results from the TEOAE predicted the outcome of HI among children that had previously passed NHS. This implies that some ears likely had a HI that could have been identified at birth, if stricter TEOAE passing criteria were in place.

Given the success of NHS, it may be easy to neglect the fact that not all infants with congenital HI are detected by NHS (Joint Committee on Infant Hearing, 2019). Study V corroborated early findings from Norton et al. (2000) showing evidence of the existence of false negatives in an NHS programme. However, it is impossible to estimate the actual false negative rate in the retrospective design of study V because of the likelihood of delayed-onset HI in the study group.

In study V, 59% of ears with HI that passed TEOAE had mild HI. Traditionally, NHS was not designed to detect infants with mild HI (Joint Committee on Infant Hearing, 2019). This is evidenced by results of study I. Across NHS programmes surveyed, the target HI for NHS was over 35-40 dB HL for 23 programmes. For an additional nine programmes, the target HI included mild HI of less than 35 dB HL. There is growing evidence for the importance of intervention for children with mild HI (Walker et al., 2015), though uncertainty remains regarding if and how children with mild HI should be intervened (Ching et al., 2021). As more evidence becomes available on best practice for infants with mild HI, modifications to the NHS protocol to improve false negative rates may be warranted so that these infants are detected earlier, and appropriate and timely interventions can be offered.

5.4 EVALUATION OF PROGRAMMES AND PROTOCOLS FOR NEWBORN HEARING SCREENING (Study I, II, IV and V)

Results from studies I, II, IV and V described protocol and programme components of NHS and the respective outcomes. Various mechanisms in an NHS programme were connected to potential outcomes, which may be modified by contextual factors. Some key components and processes of an NHS programme are discussed in the following sections.

5.4.1 Protocol

The protocol used for screening was a variable of interest for studies I, II, IV and V. The protocol included the test method used for screening, the number of steps, and the passing criteria.

5.4.1.1 Using aABR for high-risk infants

For the countries or regions surveyed in study I, the test method and number of steps used for high- and low-risk infants are described in the Appendix (Table A-1). As shown in Figure 7, countries included in the cluster *single-protocol for all infants with OAE-only* had significantly lower health expenditure per capita and human development index than countries in the cluster *dual-protocol with aABR included for low-risk infants*. These results suggest that countries with lower health spending chose not to include aABR technology, even for high-risk infants. It is well recognized that aABR should be used for infants in the NICU (EFCNI et al., 2018; Joint Committee on Infant Hearing, 2019), due to the increased prevalence of ANSD among this population compared to well babies (Berg et al., 2005; Rance et al., 1999). The start-up costs for aABR equipment are estimated to be about double the cost of OAE (Boshuizen et al., 2001; Lemons et al., 2002; Ong et al., 2020; Vohr et al., 2001). This may deter programmes with limited resources to purchase aABR devices for high-risk infants, opting instead for an OAE-based programme that reaches all infants.

5.4.1.2 Using aABR for low-risk infants

The implications of using aABR for low-risk infants were investigated in studies I, II and IV. Results from study II showed that NHS programmes that used aABR in their low-risk screening protocol had overall lower referral rates to diagnostic assessment (pooled referral rate of 1.7%) compared to protocols using OAE only (pooled referral rate of 2.1%). Study IV showed that if a two-technology screen was used before discharge (first OAE and then aABR if OAE fails), the overall referral rate was lower, compared to using OAE only (Gravel et al., 2000; Lin et al., 2005; Lin et al., 2007; Shang et al., 2016). Results from study I showed that countries in the cluster that used aABR for low risk infants had the highest health expenditure per capita and highest human development index. However, using aABR may be a worthwhile investment even for countries with limited resources, given the cost-savings from lower referral rates. Verkleij et al. (2021) demonstrated that a two-technology protocol, OAE and aABR before discharge, is the most cost-effective option for well-baby screening in Albania.

From study I, only two NHS programmes used aABR in step 1 for low-risk infants. In study IV, 14 reports were synthesized that described the differences in referral rate between TEOAE and aABR as the initial screening method for low-risk infants or all infants combined. Twelve of 14 studies reported lower referral rates from step 1 using aABR compared to TEOAE. Eight studies showed that the referral rate with aABR was 1% to 9% lower than with TEOAE. Three studies showed a much greater reduction in referral rate, with differences of 36%, 44% and 64%.

Using aABR as the primary screening method for low-risk infants is cost-effective according to an economic evaluation in Australia where aABR is used (Sharma et al., 2022). An earlier study by Lemons et al. (2002) demonstrated that aABR is an overall cheaper option compared to TEOAE due to the lower referral rate achieved with aABR, despite the difference in initial technology costs. In contrast, Verkleij et al. (2021) reported that the aABR-only protocol was more expensive than protocols using an OAE-only or a two-technology protocol; however, it also resulted in a more effective programme given the lower referral rate and lower risk that infants are missed between multiple inpatient screens. Therefore, the decision to use aABR instead of OAE for step 1 may depend on the contextual factors. For example, aABR may be preferred over OAE if LTFU rates are particularly high. Countries or regions with early discharge from the maternity ward may also benefit from an aABR-only protocol, especially if there is a low participation rate to postnatal follow-up appointments. NHS programmes that are struggling with high referral rates from OAE screening may choose to pilot an aABR protocol to attempt to reduce referral rates. On the other hand, study IV showed a wide range of referral rates for both OAE (3% to 71%) and aABR (1% to 23%) across the 14 studies synthesized. This demonstrates that technology choice is not the only reason for variation in referral rates. Therefore, quality improvement may begin by brainstorming the specific reasons behind the high referral rate, prior to adjusting the protocol.

5.4.1.3 Number of steps or screens

The number of steps was investigated in studies I and II. The most common low-risk protocols among the NHS programmes surveyed were a 2-step OAE protocol and a 3-step OAE, OAE, aABR protocol. Figure 9 illustrates the impact of multiple screening steps on the referral rate, extracted from study II. The pooled final referral rates for protocols with aABR were 1.66% for programmes with a 2-step protocol and 0.80% for programmes with a 3- or 4-step protocol. Although more screening steps can improve the false positive rate, it also increases the risk of LTFU between steps. This reduces the overall effectiveness of the screening programme (Verkleij et al., 2021). Therefore, an analysis of context specific to the country or region where screening is implemented may be a beneficial first step toward deciding the number of steps and the test method in the protocol.

The number of *screens* is defined here as the number of repeat attempts to achieve a pass within a single step. In study IV, the use of multiple screens in step 1 was a variable measured in six reports. For both OAE and aABR screening, performing a repeat screen

before discharge from the maternity ward lowered the referral rate. For some reports, these two screening attempts occurred hours apart and thus could also be related to the age of the infant when tested (Burdzgla et al., 2007; Shoup et al., 2005). However, for two other reports, rescreening was performed immediately after a fail with OAE, still resulting in a reduction of the referral rate from 5.2% to 3.6% in one report (Korres et al., 2005) and from 15.4% to 8.7% in another report (Vernier et al., 2021). Some example reasons for such a reduction in false positives from rescreening could be a blocked probe that is replaced or a noisy infant that is settled.

Given the inconsistency in the definition of a "step" described earlier, clarification questions were posed to Country Representatives as to whether multiple *screens* were performed in one step. The results are denoted with an asterisk in Table A-1 (appendix). Most Country Representatives could not report on whether multiple screening attempts were performed in their protocol, which implies that some screeners may naturally attempt a rescreen while others do not. Conversely, with too many screening attempts, the risk of a false negative increases (Hyde, 2016; Joint Committee on Infant Hearing, 2019). Therefore, standardization of such procedures could provide improvements in the false positive rates while also minimizing the risk of false negatives due to over-screening.

5.4.1.4 Passing criteria

Passing criteria were analyzed in studies I, IV and V. Passing criteria were surveyed in study I. In study IV, reports that investigated the impact of passing criteria on referral rate were synthesized. In study V the passing criteria used for OAE were investigated more thoroughly as they relate to the likelihood of false negatives.

In the EUSCREEN survey, out of the 26 NHS programmes that used OAE, 20 reported that passing criteria were defined by the screening device. For six programmes, Country Representatives reported on the criteria used to define a passing OAE test (DPOAE or TEOAE). These criteria ranged from 3 to 6 dB SNR at three to five frequency bands. In Stockholm Region, a 70% reproducibility criterion was also applied. For aABR, passing criteria were defined in terms of the minimum stimulus level where a response is detected. This level ranged from 30 to 45 dB nHL across 25 NHS programmes that used aABR for high- and/or low-risk infants. Out of these, 18 used 35 dB nHL as the passing intensity.

In study V, TEOAE data were analyzed for suspected false negatives and matched to a group of controls with normal hearing. Based on the findings, implications can be made with regards to false negatives and passing criteria. The results from study V showed the influence of two variables in final the logistic regression model, described in

Table 8. These findings suggest that, if an SNR of 6 dB was used at 4 kHz, or if a minimum TEOAE amplitude in the mid-frequencies was applied, then more infants with HI would have failed OAE screening. The majority of these infants had mild HI.

Table 8. Results of the conditional logistic regression analysis showing two significant variables from the optimal model. The outcome 1 was hearing impairment (HI) and 0 was no HI.

Significant variables in final model	p-value	Odds ratio	95% confidence interval	Interpretation of odds ratios
SNR of ≥ 6 dB at 4 kHz (ref: < 6 dB)	<0.001	0.24	0.11 to 0.51	The odds of a HI decreases by a factor of 4 when SNR ≥6 dB.
TEOAE amplitude, mid-frequencies (1.5 and 2 kHz)	<0.001	0.90	0.87 to 0.93	The odds of a HI decreases by 10% for each 1 dB increase in TEOAE amplitude

The potential consequence of increasing the stringency of the passing criteria is the risk for increasing false positives. In study V, 2.8% of ears without HI would have theoretically failed had 6 dB SNR been a required criterion for a pass. However, in study IV, four reports compared referral rate across passing criteria for OAE in a field setting (De Ceulaer et al., 1999; Gabbard et al., 1999; Korres, Balatsouras, et al., 2003; Korres et al., 2005). All reports showed negligible to no differences in referral rate after increasing the passing criteria from 3 to 6 dB SNR across the required frequency bands. Therefore, the benefit of detecting more children with HI may outweigh the risk of more false positives.

For aABR, modifying the passing criteria means altering the minimum stimulus intensity. Although aABR was not evaluated with regards to false negatives in this thesis, it has been investigated by others. Johnson et al. (2005) showed that approximately 23% of infants with HI passed aABR at 35 dB nHL, and Levit et al. (2015) found that 52% of infants with HI passed aABR at 45 dB nHL. The majority of infants that passed aABR in both studies had mild HI. The decision to increase stringency for passing criteria may rely on the situation in the country or region. Example contextual questions are, what is the existing pathway for intervention for children with mild HI, what is the current demand for intervention, what is the healthcare infrastructure and attitudes of clinicians for intervening mild HI among infants and young children, what is the existing false positive rate, and what are the wait times for diagnostic assessment among children who fail screening.

5.4.2 Age of the infant at initial screen

Studies I, II, and IV assessed the age that low-risk infants should complete step 1. For study I, the question was categorical: <24 hours, 24-72 hours, or >72 hours, though more than one category could be selected. Six programmes exclusively screened infants after 72 hours from birth. One programme exclusively screened infants before 24 hours from birth. The remaining 32 programmes screened between 24-72 hours from birth; three of these programmes also screened after 72 hours and two also screened before 24 hours from birth.

In study II, the reported referral rates were compared to the minimum age at screening. The referral rate from step 1 was reported by Country Representatives from 13 NHS programmes. For two programmes that screened infants prior to 24 hours from birth, referral rates were 6% and 22%. The pooled referral rate for the eight programmes that screened from 24 hours

onward was 7.5% (range of 1.8% to 15.3%), and for three programmes that screened from 72 hours onward, referral rates were 3.8%, 4.0% and 4.3%.

The impact of age on referral rate is a common topic in the literature, as screening at the optimal age can reduce the risk of false positives due to the presence of fluid or debris in the ear after birth (Kemp & Ryan, 1991). In study IV, eleven reports compared the referral rates from step 1 across the first week of life, for low-risk infants or all infants combined. For aABR, one report showed a decline from 22% to 11% within the first two days (Kelly et al., 2021), and the other showed no noticeable difference, though referral rates were around 1% from day 1 (Chung et al., 2019). For OAE, all nine reports showed decreasing referral rates from birth up to 3 days of age (Arslan et al., 2013; Berninger & Westling, 2011; Dimitriou et al., 2016; Hrncic et al., 2019; Korres, Nikolopoulos, et al., 2003; Tabrizi et al., 2017; Vernier et al., 2021; Vohr et al., 1993; Wessex Universal Neonatal Hearing Screening Trial Group, 1998).

After day 3, the trends differed across reports. In some reports, a plateau was reached around day 3 to 4 onward (Arslan et al., 2013; Tabrizi et al., 2017; Vohr et al., 1993). In three other reports, referral rates increased slightly after day 4 (Berninger & Westling, 2011; Dimitriou et al., 2016; Wessex Universal Neonatal Hearing Screening Trial Group, 1998). Two studies also showed higher referral rates when screening was performed 2 weeks to 2 months of age, compared to a few days after birth (Kolski et al., 2007; Uilenburg et al., 2009). One possible explanation of increasing referral rates after 3 to 4 days of age is increasing rates of OME (Berninger & Westling, 2011). Another explanation might be the increasing alertness, movement, and noise produced by newborns within the first weeks of life.

From these findings, screening between 3 to 5 days of age results in the lowest referral rate. However, the optimal age may vary based on the context. For example, for countries and regions with discharge from the maternity ward before 3 days postpartum, the initial screening step will either occur earlier than 3 days in the maternity ward or during an outpatient appointment. In this case, the consequence of increasing false positives from an early screen should be weighed against the feasibility of performing an outpatient screen and the risk of poor coverage rates if existing participation rates to postnatal appointments are low.

5.4.3 Location of screening

Location was evaluated in studies I and IV. Country Representatives in study I reported that screening was performed in the maternity hospital or NICU, with the exception of three programmes.

Study IV synthesized reports that investigated the relationship between screening location for step 2 and LTFU rate. According to five reports, the location of step 2 can be an important factor determining whether families return to their follow-up screening appointment (Barker et al., 2013; Cunningham et al., 2018; Hunter et al., 2016; Thomson & Yoshinaga-Itano, 2018; Uilenburg et al., 2009). The ideal location for step 2 may depend largely on the

context. One report showed a higher LTFU when step 2 was performed in a well-baby clinic compared to a home visit (Uilenburg et al., 2009). Another report showed better follow-up among low-income families if step 2 was performed in collaboration with community infant health clinics and held at their locations (Hunter et al., 2016).

5.4.4 Screening professionals and experience

There were a variety of professionals who perform screening. The professions named by Country Representatives in study I included nurses, midwives, audiologists, and physicians (ENTs, paediatricians), other health care workers including technicians, and dedicated lay screeners. In study IV, the screening profession itself was not related to the referral rate for aABR, according to one report (Stewart et al., 2000). There was also inconsistency in the correlation between screener experience on referral rate. One report showed that the experience of the screening staff is an influential factor for referral rates with OAE but not with aABR (de Kock et al., 2016). Another study found no differences in referral rate from OAE screening across screener experience level (Gallus et al., 2020). Although excluded from study IV due to quality, Lemons et al. (2002) reported a lower referral rate due to screener experience for aABR but not OAE. In sum, there is little consistency, and other contextual factors are likely involved in the relationship between screener experience and referral rate.

For example, results from study IV showed that hospitals with higher birth rates tended to have overall lower referral rates and lower LTFU (Mehl & Thomson, 2002; Prince et al., 2003; Thomson & Yoshinaga-Itano, 2018). However, this was not consistent across reports. In one report, the high birth rate caused extra pressure on screeners to complete screening together with their other duties, causing a higher referral rate and higher LFTU than a hospital with low birth rate (Scheepers et al., 2014). In other reports, NHS was embedded in a larger regional programme, so that the smaller organization could be supported by a larger network (Fan et al., 2010; Hergils, 2007). Two additional reports also showed that amalgamation of smaller programmes resulted in better LTFU rates than if local screening sites were managed independently (Barker et al., 2013; Park et al., 2020).

The importance of audiology involvement in the screening programme to achieve low LTFU was emphasized by Thomson and Yoshinaga-Itano (2018). According to this report, if an audiologist was involved in the NHS programme, the influence of other factors on LTFU rate (i.e., hospital size, step 2 location) became no longer significant or less influential. An audiologist involved in managing the NHS programme can act as a facilitator for supporting and educating screening staff and managing the quality control aspects of the programme.

As with other mechanisms in NHS, the relationship between staffing and outcomes may depend largely on the context. These factors could include the attitudes of screening professionals, the structure and workload of existing neonatal care professionals, and funds available for hiring dedicated staff. Using existing nursing staff to carry out screening in addition to their regular duties may be a cheaper option at first (Lemons et al., 2002).

However, there is a risk that clinical staff become overburdened, resulting in lower coverage rate (Friderichs et al., 2012), higher referral rate or higher LTFU after screening (Scheepers et al., 2014). Therefore, the use of existing clinical staff for NHS could be possible if workload is managed appropriately. An audiologist on site to facilitate screening may help with quality management and improve outcomes (Thomson & Yoshinaga-Itano, 2018).

5.5 EVALUATION OF PROGRAMMES, PROTOCOLS AND OUTCOMES FOR CHILDHOOD HEARING SCREENING AFTER THE NEWBORN PERIOD (Study III)

Study III focused on childhood hearing screening after the newborn period. The protocol including the test method, other programme factors, and the available outcomes are discussed in the following sections. As described previously, no screening was reported between the newborn period and 3 years of age. This evaluation therefore focuses on preschool- and school-entry hearing screening.

5.5.1 Protocol

The test method used for preschool- or school-entry screening was predominantly pure-tone audiometry screening. In all 17 countries or regions with universal programmes, pure-tone audiometry screening was performed. Country Representatives from eight additional countries or regions reported that some screening was performed (e.g., not routine, opportunistic, or varies across the region). Of these, two reported that the whispered voice test was used, while the others used pure-tone audiometry screening. The sensitivity of the whispered voice test in children under 6 years of age ranges between 56% to 96% (Pirozzo et al., 2003; Skoloudik et al., 2020). Despite the relatively low sensitivity, no equipment is required, which makes this test a feasible option for countries or regions that lack resources for purchasing screening audiometers. No programmes reported using any alternative technology, like speech-in-noise tests.

The passing criteria for pure-tone audiometry screening varied across the 10 universal programmes where criteria were reported. All screening programmes included the frequencies 1, 2 and 4 kHz, except one which did not screen at 2 kHz. Eight programmes also screened at 500 Hz. The passing intensity ranged from 20 to 40 dB HL. Similar to NHS, deciding on the stringency of the passing intensity will depend on local factors, such as the burden on diagnostic clinics and the existing practice for diagnosing and intervening mild HI. For school-entry hearing screening, detecting and diagnosing mild HI may be more crucial, compared to the newborn stage, as studies have shown the consequences of mild HI and effects of amplification on language development in school-age children (Walker et al., 2015; Walker et al., 2020).

5.5.2 Other programme factors

The professionals performing the screening varied across the 17 programmes with universal screening. Nurses were the most common professional reported. Other professionals included speech-language therapists, audiologists, and physicians. Out of 17 programmes, the location

of screening was divided between school (nine), health clinics (seven) or both (one). Example A decision-maker may consider contextual factors when deciding on the location and screening professional, such as the existing framework for health screening in schools or well-child clinics, and the capacity for nurses or health workers to take on an additional task.

The age at which children are screened ranged from 3 to 7 years of age. Screening at the age of 3 years is advantageous for detection HI as early as possible, but it may not be suitable if there is no working system for health check-ups, or if participation rates in preschool are low at this age. The other effect of screening at 3 to 4 years of age is that more false positives are reported due to factors such as sustained attention, reliability issues with low-frequency tones, and high rates of OME (Browning, 2000). Conversely, waiting until age 6 or 7 before screening can be detrimental to children with HI who would benefit from intervention earlier.

5.5.3 Outcomes

Coverage rates were reported from three countries: 97%, 99%, and 45%. This last rate was the national coverage rate in Serbia; the regional rate in Belgrade was 92%. For referral rate, two countries reported final referral rates of 7.6% and 7.9%. Follow-up rates were poor for both countries, 58% and 77%. Similar to NHS, preschool and school-entry hearing screening suffers from issues with follow up. More detailed investigation is needed on the barriers to follow-up for these programmes.

5.6 STRENGTHS AND LIMITATIONS

5.6.1 Study I, II and III

The EUSCREEN survey presented many strengths, but also some limitations. The limitations of studies I, II and III were apparent early on in data collection. Specifically, the clarity of certain questions was poor which caused misinterpretation by the Country Representatives and issues with the validity. In order to resolve the situation, a glossary was created, and clarification questions were drafted to ensure understanding. However, it was not feasible to overburden Country Representatives with many additional questions, and therefore, selected clarification questions were posed that could offer a comprehensive overview of the status and outcomes of the hearing screening programme. In addition, the questionnaire was long, which made it difficult for many Country Representatives to remain engaged. A significant time commitment was required on part of the Country Representatives to complete all questions. The difficulty for Country Representatives to aggregate the information requested was also an unexpected obstacle. Although the Country Representatives were compensated for completing the questionnaire, these challenges could have been avoided with more planning prior to widespread release of the questionnaire, such as a more thorough piloting and translation of the questionnaire into local languages.

Country Representatives who responded to the EUSCREEN survey often worked closely with their local screening programme. Therefore, the prospect of reporting bias was not

overlooked. This limitation was previously discussed under Ethical Considerations, in addition to the steps taken to mitigate the risk.

Another limitation to the EUSCREEN questionnaire stemmed from the variability in how outcomes were defined and measured across programmes. This made it difficult to compare outcomes across programmes. For example, quality indicators were defined differently across programmes. Additionally, questions on specific outcome (e.g., the percentage of infants screened diagnosed with HI >70 dB), some Country Representatives remarked that providing this value was not possible, as other threshold values were used for reporting on the detection rate (e.g., 60 or 80 dB). This barrier for comparing data across NHS programmes will continue unless universal agreement is made for standardization of quality indicators (Alam et al., 2016; Mason et al., 2008).

Notwithstanding the limitations of the EUSCREEN questionnaire (studies, I, II and III), this study offered a comprehensive evaluation of childhood hearing screening across a large number of countries, primarily in Europe. Due to the in-depth cross-checking and validation methods, these studies resulted in a large volume of high-quality information. Procedures for validation procedures were necessary to avoid inaccuracies due to limitations described. Moreover, the variability in data quality across programmes, albeit a limitation, also became a key finding of the EUSCREEN project. It underscores the requirement for thorough documentation and quality management in EHDI.

5.6.2 Study IV

Study IV was a systematic review of the literature. The risk of reporting bias is discussed in the ethical considerations (under Methods). Two additional limitations can be drawn from this study. First, there were 900 records that met title/abstract criteria for inclusion, yet were further excluded because they were not written in English. Although the large majority of these reports would have been excluded from the full-text sorting, the language bias is concerning given the quantity of non-english records located. It is possible that NHS programmes tested various options for implementing or modifying their screening programme, which were then published in local journals in the local language. Unfortunately, the resources were not available for translating the 900 reports in 25 different languages. However, the reports that were included in the review originated from 31 countries and six continents. Therefore, despite the language bias, study IV offered a wide view of the protocol and programme factors influencing NHS programmes around the world.

Next, the issue with defining LTFU is apparent in study IV. This was discussed previously as a finding in this review; however, it is also important to consider as a limitation as it affects the generalisability of the LTFU results. As described for the EUSCREEN questionnaire, there is no standardization in how quality indicators are defined across NHS programmes. Given this variability noted, it is important for future reports on NHS outcomes to define their calculations for all indicators used.

The strength of study IV is its comprehensiveness and practicality. This study was a large investigation into the mechanisms and outcomes of NHS, covering a broad range of components that make up a screening programme. A huge body of literature was screened, evaluated, and synthesized. The results offer a foundation for EHDI decision-makers to evaluate the key options in an NHS protocol or programme factors for implementing or modifying a programmes in their context.

5.6.3 Study V

The limitation in study V was the likelihood of delayed-onset or acquired HI in the study group. Study V retrospectively evaluated TEOAE results to predict whether certain criteria could distinguish ears with HI from normal hearing. The study group was restricted to children detected with HI before age 8; however, there are multiple causes of HI in childhood after the newborn period. Medical records were not accessible in this study; therefore, it is probable that some children who acquired HI after birth were included in the study group. Furthermore, delayed-onset HI cannot be definitively separated from congenital HI even if medical records were accessed. This remains a limitation in all research that evaluates childhood HI after the newborn period. This bias would have caused an underestimation of the effect. However, even with this underestimation, the results of study V showed that certain TEOAE criteria significantly differentiated the two groups.

Despite its limitation, study V provided new evidence that NHS programmes miss children with HI. Outcomes from NHS programmes typically include the coverage rate, referral rate, follow-up rate and detection rate (Joint Committee on Infant Hearing, 2007; Wood et al., 2015). The rate of children with HI missed by screening is an outcome that is widely overlooked (Norton, Gorga, Widen, Folsom, et al., 2000), likely due to the difficulty in distinguishing these children from delayed-onset HI, as previously discussed. Study V utilized the unique registry in Region Stockholm storing TEOAE data for all children born since 2006. Because of the availability of these data, the results of study V could evaluate the presence of false negatives, something that is rarely possible in NHS programmes, and a novel evaluation of TEOAE variables and human error in NHS data management.

6 CONCLUSIONS

The following conclusions can be drawn.

- In many high-income countries with high health spending, universal newborn hearing screening programmes exist, and best practice guidelines are followed for high-risk infants. However, progress still needs to be made in more resource-constrained countries to improve the detection of HI for all infants.

- Preschool or school-entry hearing screening programmes are not widely available. There is no clear consensus on whether preschool/school-entry hearing screening should be implemented, sustained or discontinued.

- Lost to follow-up is a ubiqutous problem for childhood hearing screening programmes. Furthermore, some children with HI who fail screening may be incorrectly documented as having passed. These issues present a major concern that children with HI are not diagnosed early due to LTFU or errors in documentation.

- By altering the passing criteria for TEOAE (i.e., using 6 dB SNR at 4 kHz and a minimum TEOAE amplitude in the mid-frequencies), the detection of HI can be improved. Increasing the stringency is unlikely to significantly affect the referral rate.

- Some NHS programmes do not use aABR for high-risk infants, despite widely recognized recommendations. These countries have, on average, lower health care expenditure per capita than countries that follow recommendations. The benefit of universal screening with a cheaper technology may outweigh the benefits of detecting cases of ANSD.

- NHS programmes that include aABR in their protocol for low-risk infants had the higher health expenditure per capita, on average. Because aABR reduces the referral rate to diagnostic assessment, it may be a cost-saving addition to protocol. This may also be true for using aABR for step 1. However, there was a wide range of referral rates across programmes that used either screening method. Therefore, the decision to replace OAE with aABR for step 1 should be made with consideration of the local quality improvement needs of the NHS programme.

- Factors that might be considered to reduce the referral rate are: the number of steps in the protocol, repeat screening within a step, age of the infant at the initial screen, the screener experience and the organisational structure of the NHS programme (e.g., smaller NHS programmes networking with larger programmes).

- Factors that may be considered to reduce the LTFU from screening step 1 are: the location of step 2 screening, the experience of the screeners, and having an audiologist involved in the programme.

- Data availability was poor. Just over half of the participating NHS programmes had valid referral rates, and less than a third had valid rates for follow-up from screening. For childhood hearing screening programmes after the newborn period, only two could provide valid data on these performance measures. In order to apply any of the other findings from this thesis, the initial step for any childhood hearing screening programme is to assess its existing quality. This is done by defining quality indicators, and then collecting and reporting on the key outcomes from screening, diagnosis and intervention. With a sustainable system for monitoring and evaluating outcomes over time, decision-makers can then assess the areas of concern and apply these strategies for improving their childhood hearing screening programmes.

7 POINTS OF PERSPECTIVE

7.1 IMPLICATIONS FOR DECISION-MAKERS

The realist evaluation framework used in this thesis provided a structure for how the mechanisms of a childhood hearing screening programme were related to outcomes and modified by context. Each childhood hearing screening programme is unique, and decision-makers must make certain choices when implementing or modifying their NHS programme. Decisions on key parameters may include targeted versus universal screening, the target condition for screening, the test methods, the passing criteria, the number of steps, the location of screening, the age of the infant/child when screened, and the organizational structure of the programme. Table 9 provides a list of examples derived from the results of this thesis that may be useful for decision-makers implementing or modifying an NHS programme. The definition of decision-makers are any group or individual who makes decisions with regards to EHDI. This could include government policymakers, EHDI expert committees, programme managers, screening professionals, and screening device manufacturers.

Some practical advice derived from this thesis is presented below:

Perform a cost-effectiveness analysis prior to implementation or de-implementation of preschool or school entry hearing screening. The results of this thesis revealed no consistency across countries or regions toward implementation or de-implementation of preschool or school-entry hearing screening. This inconsistency is mirrored in the literature on the cost-effectiveness of these programmes (Fortnum et al., 2016; Gumbie et al., 2022).

Reduce LTFU by holding rescreening appointments at an accessible location, schedule rescreening directly, allow screeners to gain experience without becoming overburdened, and have an audiologist involved in the screening programme. LTFU remains a major obstacle for many of the NHS programmes surveyed, and for the two preschool or school-entry hearing screening programmes that provided this outcome.

Be aware that low TEOAE amplitude or SNR that barely meets passing criteria may be evidence of a mild HI. Increase the stringency of passing criteria to detect more infants with congenital HI. The SNR at 4 kHz could increase to 6 dB, and a minimum amplitude could be set in the mid-frequencies.

Consider the risk of false negatives with setting the default passing criteria used and increase the transparency of the algorithm. It is important for EHDI decision-makers to be critical of the passing criteria used in automated devices, to ensure that referral rates are reasonably low, yet infants with HI are not being missed.

Streamline and digitalize the screening results from the device directly into a tracking database. Some children with HI should have failed the screening but were incorrectly entered into the database as a pass. An automated transfer of screening results will mitigate the risk of error.

Consider solutions to reduce false positives and write them into protocol, such as using aABR, screening 3 to 5 days after birth, increasing the capacity for screeners to gain experience and network between sites, rescreening immediately after an initial fail, and performing screening only when the infant is calm. Incorporating aABR into the screening protocol after OAE fail will reduce referral rates to diagnostic assessment. Using aABR instead of OAE as the primary screening method may reduce referral rates from step 1 and detect cases of ANSD. Screening on 3 to 5 days after birth is the optimal window for screening. Additionally, how a programme is networked between local screening sites and the experience and workload of the screening staff can influence the performance of the screening programme. Finally, other factors that may seem trivial can influence referral rate, such as rescreening immediately after a fail and screening only when the infant is calm.

While the results of this thesis offer ideas for decision-makers to modify their screening programme and improve results, decisions should only be taken based on the needs and circumstances within a certain context. Preschool- or school-entry hearing screening may take place either in a school setting or in the well-child clinic depending on context. The advantages and disadvantages of screening younger versus older children discussed in this thesis should also be considered. For NHS, the decision on who should perform screening (e.g., a dedicated lay screener or health care worker such as a nurse or midwife) should be considered relative to factors such as existing workload and capacity for training. Similarly, the existing framework in postnatal care can help in deciding whether screening is performed in the maternity hospital or in an outpatient setting.

Improve documentation and definitions in the screening protocol. There were insufficiencies in how protocols and quality indicators were documented and defined across screening programmes. In the systematic review, many studies failed to define LTFU. Furthermore, many Country Representatives could not provide details on details of the protocol, such as the passing criteria or whether screeners performed multiple immediate rescreens after a failed test. These seemingly negligible parameters were shown in this thesis to contribute significantly to screening outcomes. Therefore, they should be written into protocol.

Implement sustainable procedures for quality management. Many Country Representatives could not provide valid data on the performance of their screening programme. Although some data may be available internally, it is likely that for many countries or regions, quality management was not regularly performed. Quality management should be in place to ensure an effective EHDI programme (Joint Committee on Infant Hearing, 2019). This should include ongoing data collection, evaluation and reporting of predefined quality indicators.

Table 9. Examples of the realist evaluation for components of newborn hearing screening, including the mechanisms and outcomes with their potential influential contextual factors.

	Mechanism	**Examples of contextual factors**	**Outcomes**
Test method	Using aABR instead of OAE for step 1 before discharge from maternity hospital	- Low attendance to postnatal follow-up after maternity discharge - Policy for early discharge from maternity - High LTFU - Funding available for implementing aABR - Priority for ANSD detection among well babies	- Increased coverage rate - Lower referral rate from step 1 - Increased detection of ANSD for well babies
	Using aABR for step 2 or step 3 after previous OAE fail	- Funding available for purchasing aABR devices - Infrastructure for aABR-trained personnel (and devices) at step 2 or 3 follow-up locations - High referral rate; burden on diagnostic services	- Lower referral rate to diagnostic assessment
	Using OAE for all infants (including high-risk)	- Limited resources allocated for screening - Little experience with newborn hearing screening - Untrained screening personnel	- Higher coverage rate - Faster training and implementation - Lower costs - No detection of ANSD among high-risk infants
OAE pass criteria	Increase stringency to include SNR of 6 dB and minimum TEOAE amplitudes at select frequency bands	- Low refer rates in current screening programme - Clinical recognition for necessity of intervention of mild HI - Clinical pathway for mild HI follow-up and/or intervention	- Higher referral rate - More children with HI detected by OAE screening
Number of steps or screens	Increasing from 2 to 3 steps	- High follow-up rates from step 1 to step 2 - High referral rate from step 2; burden on diagnostic services	- Lower referral rate to diagnostic assessment - Risk for more infants LTFU
	Reattempting the screen at step 1 after initial fail	- High step 1 referral rate	- Lower referral rate from step 1 - Higher risk for false negative
Age at initial screen	Extending the age at step 1 from 1 day to 3 days after birth	- Existing framework for postnatal follow-up - High follow-up rate to postnatal appointment - High referral rate from step 1	- Lower referral rate from step 1 - Risk for reduced coverage rate

Location	Step 1 performed in maternity hospital before discharge	- Maternity ward stay is at least 48 to 72 hours - Poor infrastructure for postnatal follow-up with high returns rates	- Higher coverage rate - Higher referral rate from step 1
	Step 2 performed in familiar, accessible location	- Geographical or economic conditions among families hinder travel - Maternity hospitals are accessible to all families	- Lower LTFU between steps 1 and 2
Screeners and experience	Employing a dedicated screener or technician without medical or audiological background	- Funding and programmes allocated for intensive training - Management is experienced personnel with audiology knowledge - Other health care professionals (i.e., nurses) would be overburdened if taking on additional workload	- Lower referral rates - Lower LTFU
	Employing an audiologist as a facilitator and quality manager for screening	- Funding available for the position - Attitudes of screening staff and other neonatal care workers	- Lower LTFU
	Amalgamate smaller programmes to a larger network to increase quality management, standardize performance indicators, and improve knowledge exchange	- Existing network for quality management and regulation across local sites - Funding allocation in place - Knowledgeable management in larger centre	- Decreased referral rate - Decreased LTFU

7.2 REMAINING GAPS AND FUTURE DIRECTIONS FOR RESEARCH

This thesis raises many further questions and theories regarding the effectiveness of EHDI and strategies for screening. The following section discusses some of the directions for future research.

First, this thesis provided an evaluation of childhood hearing screening programmes using quantitative measures. Notably, qualitative measures were missing from this thesis, which can offer insights that were not possible with a quantitative-only design. For example, interviews with Country Representatives could uncover the aspects of the screening programmes that do or do not function well in their context, which could explain why certain protocol decisions were chosen in various countries or regions.

Moreover, qualitative interviews could unravel the specific reasons for the lack of valid outcome data. It is established that a system for programme evaluation is needed for all EHDI programmes (Joint Committee on Infant Hearing, 2019) and a prerequisite for screening (Andermann et al., 2008); however, this thesis showed that many Country Representatives could not provide information on the performance of their screening programme. What are the reasons that EHDI programmes do not have effective quality management? What are the barriers for implementing and sustaining a system for monitoring its performance? These questions can be best addressed using qualitative interviews.

This thesis also lacked a multi-stakeholder perspective, which is a necessary future direction in evaluation studies of EHDI programmes. In this thesis, an ecological study was performed aggregating information from experts in the field. This approach is commonly used to assess the status of programmes across many countries or regions (Bagatto et al., 2020; Bubbico et al., 2013). However, decision-making in health care also should account for patient-level concerns (Clayman et al., 2015). Future investigations into the decision models for EHDI in various contexts should therefore include a range of stakeholders, specifically families, screening professionals, and individuals with HI.

Likewise, the data aggregated in this thesis represented the NHS programme as a whole and did not divulge any health disparities, barriers to access, or other injustices in EHDI care. For example, reporting a coverage rate of 95% may meet benchmarks, but does not describe whether the 5% of infants who are not covered by screening are at a disadvantage due to their sociocultural, regional, or economic situation. The next steps for evaluating the performance of EHDI programmes should therefore consider the health disparities within each programme.

One issue revealed by the results of this thesis is LTFU. It is revealed that LTFU remains a ubiquitous problem for NHS programmes, and factors were studied that influenced LTFU on an organizational level. Other studies have shown that family or maternal-level factors also impacted the risk of LTFU (Cunningham et al., 2018). However, there remains a lack of evidence regarding the strategies that are effective at overcoming these barriers to follow-up. This is a necessary future direction in EHDI research.

Finally, the question remains, what is the most effective strategy for detecting HI after the newborn period? This thesis focused specifically on the strategies for screening and did not evaluate other methods for detection, such as speech and language milestone check-ups or the surveillance of children with risk factors. The results of this thesis showed variability in the provision of preschool or school-entry hearing screening. A study comparing the effectiveness of methods after newborn screening is required to uncover the optimal strategies for detecting all children with HI.

8 ACKNOWLEDGEMENTS

To *Inger Uhlén*, my main supervisor. Thank you for your wisdom, your patience, your critiques, and remarks, your guidance, and for your strong will to see me succeed. Your passion for helping children with hearing loss and your ambition to continue to move this field forward is an inspiration. You provided all the right ingredients for me to carry out my PhD in a supportive and enriching environment and to propel my future research career. Thank you for all that you have done!

To *Elina Mäki-Torkko*, my co-supervisor. You have taught me so much about the details of research, and particularly with regards to improving clarity in my writing. Your insight into performing systematic reviews was extremely valuable over the *multiple* years it took to carry out that study. I am particularly in debt to your comments and feedback on how to effectively write grant applications – a skill that I no doubt will continue to develop. Your positive words regarding our research will stay with me for a long time.

My colleague and friend, *Andrea Bussé.* What would this PhD be without you? My collaboration with you over the four years of the EUSCREEN project was a journey to say the least! I am so proud of the accomplishments we have made together, and while it was often not easy, I believe that just because I was able to work with you on this project I have become a stronger and better researcher on all fronts. Thank you for supporting me, for inspiring me, for challenging me, and for just being there for me. We did it!

To all partners and collaborators of the EUSCREEN hearing group, and particularly to *Hans Hoeve, André Goedegebure, Gwen Carr, Adrian Davis, Herb Simonsz and Birkena Qirjazi*. Thank you for the countless discussions and email conversations, which challenged my thinking about childhood hearing screening. To be able to work with and learn from such experts in the field has been humbling and a true honour!

To *Jill Carlton, Helen Griffiths and Paolo Mazzone*. Thank you for your collaboration and discussions, particularly early in the project, to brainstorm together from different fields how to overcome our common challenges.

To *Eveline Heijnsdijk* and *Mirjam Verkleij*. I am so grateful for everything I have learned from you regarding cost effectiveness. Not only has it been enlightening for the EUSCREEN project, but it has also given me inspiration for possible future directions in research. I hope we will be able to collaborate again!

To all the Country Representatives representing hearing screening programmes in the EUSCREEN project. *B. Qirjazi, D. Holzinger, L. Stappaert, B. Vos, F. Brkić, P. Rouev, X. Peng, M. Velepic, C. Thodi, J. Drsata, T. Ovesen, M. Bambus, M. Lepplaan, B. Ellefsen, R. Niemensivu, T. Willberg, F. Denoyelle, P. Matulat, T. Nikolopoulos, A. Gáborján, I. Hinriksdóttir, Z. Chaudhurri, G. Norman, L. Rubin, A. Martini, D. Spanca, M. Audere, S. Kušķe, N. Drazdiene, E. Lesinskas, J. M. Hild, M. Cakar, T. Fenech, W. Mulwafu, D.*

Chiaburu, T. Kujundžić, E. Zvrko, A. Meuwese, A. Goedegebure, H. Hoeve, V. Nagaraj, G. Greczka, L. Monteiro, M. Georgescu, G. Tavartkiladze, L. Gouma, S. Filipovic, G. Jokovic, L. Langova, I. Sebova, S. Battelino, D. W. Swanepoel, F. Núñez-Batalla, J. M. Sequi-Canet, I. Uhlén, B. Nora, M. Baydan and J. McCall. This project was not possible without your hard work and the countless hours you put in to completing the questionnaire. It has been an honour to be part of a project that brings together so many experts from so many different countries. Thank you for your patience with filling out all clarification questions, and for all the greetings and discussions we had at various conferences throughout the project. I hope that we will have the opportunity to see each other again.

To *Åsa Skjönsberg*, my mentor. Thank you for all your advice over the years, our meetings, emails, and general check-ins. It was truly comforting to know that I could always come to you when needed. I am looking forward to working more with you in the future!

To *Anna Persson, Marlin Johansson,* and *Satu Turunen-Taheri*. It was wonderful knowing that each of you are (or were) traveling along this same road. A simple message from one of you about an update to an article or a stipend granted would lift me up for days. Thank you for all your kindness and words of support throughout this journey. I am certain that it's only the beginning.

To *Martin Eklöf* and *Magnus Vestin*. A big thanks for your help with data aggregation from registries and databases in my last study!

To the staff at Barnhörselhabilitering at Rosenlund. Thank you for allowing me a space to carry out my PhD over the last years! We saw each other for stretches between my multiple parental leaves and working from home during the pandemic. But having a familiar place to sit and work in peace and quiet has been a true gift. It has been a real pleasure to see your smiling faces! A special thanks to *Mårten Westermark* for all the technical support you have provided over the years!

To *Leif Hergils, Karin Stenfeldt* and *Kristi Sidney Annerstedt*. A special thanks to you, the committee of my half-time review, for your insightful remarks and challenging questions during my half-time seminar. Thank you for taking the time to read carefully through my half-time report and for your encouragement to complete the final two years of my PhD.

To *Agneta Wittlock,* the all-knowing. Thank you for answering my countless questions about all things administrative, and to *Lars Olaf Cardell*. Thank you for all your support!

Thank you to those who have inspired me to purse my PhD, including my former colleagues at Vivosonic, particularly to *Aaron Steinman* for your continued support.

To the memory of *Susan Small*, my master's thesis supervisor, who encouraged me to purse this path in research. You believed in me since the beginning, opened so many doors for me,

and taught me the fundamentals of research. Moreover, you taught what it means to be a mentor. What will the field of pediatric audiology be without you? You are sorely missed!

To my parents, *Bill and Karen Mackey*, if there is anyone who carried me the furthest to reach this moment, it is the two of you. You have been my number one fans since day 1. Thank you for everything you have done to help me reach this point.

To my husband, *Gustav Hagerling*, amongst all the struggles and stress these past years you are always solid like a rock. I am so grateful that we both have passion for our careers, and that we can support each other through our work because we understand that passion. Thank you for supporting me through this milestone and I promise I will support you through yours.

To my brothers and their partners, *Jeff and Scott Mackey, Leah Diedrich and Laia Meghdadi*. Thank you for your support throughout this time, and particularly for being the best uncles and aunts my kids could ask for.

To all my in-laws, *Stephan and Mary Hagerling, Carl and Catarina Hagerling, Oscar Esping* and *Lotta Dufvenius, Lotti Wennersten, Joakim Johansson, Emma and Martin Johansson, Daniel Bengtsson, Emily Jönsson*, and *Isabelle and Anton Granqvist*. Thank you for accepting me into your family and for your help navigating all things Swedish.

To my dear friends in Canada and in Sweden who has stood beside me during this journey. Thank you for being there with a listening ear, a shoulder to lean on, and a smile and laugh to take my stress away.

Finally, to my wonderful children, *Nova and Elise*. I am so grateful that, before earning the PhD title, I earned the title "mamma". You bring joy into each and every day of my life. You are my everything.

The work in this project was funded by the European Union's Horizon 2020 research and innovation program under grant agreement no. 733352, Hörselskadades Riksförbund Hörselforskningsfonden, Stiftelsen Sunnerdahls Handikappfond and Stiftelsen Tysta Skolan.

9 REFERENCES

Alam, S., Chung, W., Deng, X., O'Hollearn, T., Beavers, J., Cunningham, R. F., . . . Do, T. N. (2018). Restructuring Data Reported from Jurisdictional Early Hearing Detection and Intervention (EHDI) Programs: A Pilot Study. *Journal of early hearing detection and intervention*, *3*(1), 57-66.

Alam, S., Satterfield, A., Mason, C. A., & Deng, X. (2016). Progress in Standardization of Reporting and Analysis of Data from Early Hearing Detection and Intervention (EHDI) Programs. *Journal of early hearing detection and intervention*, *1*(2), 2-7.

Andermann, A., Blancquaert, I., Beauchamp, S., & Dery, V. (2008). Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. *Bull World Health Organ*, *86*(4), 317-319.

Arslan, S., Isik, A. U., Imamoglu, M., Topbas, M., Aslan, Y., & Ural, A. (2013). Universal newborn hearing screening; automated transient evoked otoacoustic emissions. *B-ENT*, *9*(2), 122-131.

Bagatto, M., Moodie, S., Fitzpatrick, E., Kealey, C., Campbell, B., & Aiken, S. (2020). Status of Early Hearing Detection and Intervention Programs in Canada: Results From a Country-Wide Survey. *Canadian Journal of Speech-Language Pathology and Audiology*, *44*(3).

Barker, M. J., Hughes, E. K., & Wake, M. (2013). NICU-only versus universal screening for newborn hearing loss: Population audit. *Journal of Paediatrics & Child Health*, *49*(1), E74-79.

Barreira-Nielsen, C., Fitzpatrick, E., Hashem, S., Whittingham, J., Barrowman, N., & Aglipay, M. (2016). Progressive Hearing Loss in Early Childhood. *Ear and Hearing*, *37*(5), e311-321.

Berg, A. L., Prieve, B. A., Serpanos, Y. C., & Wheaton, M. A. (2011). Hearing Screening in a Well-Infant Nursery: Profile of Automated ABR-Fail/OAE-Pass. *Pediatrics*, *127*(2), 269.

Berg, A. L., Spitzer, J. B., Towers, H. M., Bartosiewicz, C., & Diamond, B. E. (2005). Newborn hearing screening in the NICU: profile of failed auditory brainstem response/passed otoacoustic emission. *Pediatrics*, *116*(4), 933-938.

Berninger, E., & Westling, B. (2011). Outcome of a universal newborn hearing-screening programme based on multiple transient-evoked otoacoustic emissions and clinical brainstem response audiometry. *Acta Oto-Laryngologica*, *131*(7), 728-739.

Beswick, R., Driscoll, C., & Kei, J. (2012). Monitoring for Postnatal Hearing Loss Using Risk Factors: A Systematic Literature Review. *Ear and Hearing*, *33*(6).

Beswick, R., Driscoll, C., Kei, J., & Glennon, S. (2012). Targeted surveillance for postnatal hearing loss: A program evaluation. *International journal of pediatric otorhinolaryngology*, *76*(7), 1046-1056.

Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., . . . Tooher, R. (2001). Relationships Among Speech Perception, Production, Language, Hearing Loss, and Age in Children With Impaired Hearing. *Journal of Speech, Language, and Hearing Research*, *44*(2), 264-285.

Boshuizen, H. C., van der Lem, G. J., Kauffman-de Boer, M. A., van Zanten, G. A., Oudesluys-Murphy, A. M., & Verkerk, P. H. (2001). Costs of different strategies for

neonatal hearing screening: a modelling approach. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, *85*(3), F177-F181.

Boudewyns, A., Declau, F., van den Ende, J., Hofkens, A., Dirckx, S., & Van de Heyning, P. (2016). Auditory neuropathy spectrum disorder (ANSD) in referrals from neonatal hearing screening at a well-baby clinic. *Eur J Pediatr*, *175*(7), 993-1000.

Bray, P., & Kemp, D. (1987). An advanced cochlear echo technique suitable for infant screening. *British Journal of Audiology*, *21*(3), 191-204.

Browning, G. G. (2000). Influence of Age, Type of Audiometry and Child's Concentration on Hearing Thresholds. *British Journal of Audiology*, *34*(4), 231-240.

Bubbico, L., Tognola, G., & Grandori, F. (2013). Coverage and screening protocols in UNHS programmes in Italy in 2011: A nationwide survey. *Hearing, Balance & Communication*, *11*(3), 100-103.

Burdzgla, I., Pietsch, M., Chkhartishvili, B., & Kevanishvili, Z. (2007). The proper time for hearing screening in newborns. *Georgian Medical News*(144), 24-27.

Bussé, A. M. L., Hoeve, L. J., Nasserinejad, K., Mackey, A. R., Simonsz, H. J., & Goedegebure, A. (2020). Prevalence of permanent neonatal hearing impairment: Systematic review and Bayesian meta-analysis. *International Journal of Audiology*, *59*(6), 475-485.

Butcher, E., Dezateux, C., Cortina-Borja, M., & Knowles, R. L. (2019). Prevalence of permanent childhood hearing loss detected at the universal newborn hearing screen: Systematic review and meta-analysis. *PloS one*, *14*(7).

Cadena, A. C., Lindholm, N., & Stenfeldt, K. (2021). School-based hearing screening in Sweden – An evaluation of current practices. *International journal of pediatric otorhinolaryngology*, *150*, 110938.

Cheung, A., Chen, T., Rivero, R., Hartman-Joshi, K., Cohen, M. B., & Levi, J. R. (2022). Assessing Loss to Follow-up After Newborn Hearing Screening in the Neonatal Intensive Care Unit: Sociodemographic Factors That Affect Completion of Initial Audiological Evaluation. *Ear and Hearing*, *43*(2).

Ching, T. Y., Oong, R., & Van Wanrooy, E. (2006). The ages of intervention in regions with and without universal newborn hearing screening and prevalence of childhood hearing impairment in Australia. *The Australian and New Zealand Journal of Audiology*, *28*(2), 137.

Ching, T. Y. C., Crowe, K., Martin, V., Day, J., Mahler, N., Youn, S., . . . Orsini, J. (2010). Language development and everyday functioning of children with hearing loss assessed at 3 years of age. *International Journal of Speech-Language Pathology*, *12*(2), 124-131.

Ching, T. Y. C., Saetre-Turner, M., Marnane, V., Scarinci, N., Choik, C., Tulloch, K., & Sung, V. (2021). Audiologists' perspectives on management of mild bilateral hearing loss in infants and young children. *International Journal of Audiology*, 1-9.

Chung, Y. S., Oh, S. H., & Park, S. K. (2019). Referral rates for newborn hearing screening based on the test time. *International journal of pediatric otorhinolaryngology*, *127*, 109664.

Clayman, M. L., Bylund, C. L., Chewning, B., & Makoul, G. (2015). The Impact of Patient Participation in Health Decisions Within Medical Encounters: A Systematic Review. *Medical Decision Making*, *36*(4), 427-452.

Clemens, C. J., & Davis, S. A. (2001). Minimizing False-Positives in Universal Newborn Hearing Screening: A Simple Solution. *Pediatrics*, *107*(3), e29-e29.

Committee on Fetus and Newborn. (1971). JOINT STATEMENT ON NEONATAL SCREENING FOR HEARING IMPAIRMENT. *Pediatrics*, *47*(6), 1085-1085.

Cunningham, M., Thomson, V., McKiever, E., Dickinson, L. M., Furniss, A., & Allison, M. A. (2018). Infant, Maternal, and Hospital Factors' Role in Loss to Follow-up After Failed Newborn Hearing Screening. *Academic pediatrics*, *18*(2), 188-195.

Dalkin, S. M., Greenhalgh, J., Jones, D., Cunningham, B., & Lhussier, M. (2015). What's in a mechanism? Development of a key concept in realist evaluation. *Implementation Science*, *10*(1), 49.

Davis, A., Bamford, J., & Stevens, J. (2001). Performance of neonatal and infant hearing screens: sensitivity and specificity. *British Journal of Audiology*, *35*(1), 3-15.

Davis, A., & Davis, K. A. S. (2016). Descriptive Epidemiology of Childhood Hearing Impairment. In A. M. Tharpe & R. Seewald (Eds.), *Comprehensive Handbook of Pediatric Audiology, Second Edition* (pp. 89-131). Plural Publishing.

De Ceulaer, G., Daemers, K., Van Driessche, K., Marien, S., Somers, T., Offeciers, F. E., & Govaerts, P. J. (1999). Neonatal hearing screening with transient evoked otoacoustic emissions: a learning curve. *Audiology : official organ of the International Society of Audiology*, *38*(6), 296-302.

de Kock, T., Swanepoel, D., & Hall, J. W., 3rd. (2016). Newborn hearing screening at a community-based obstetric unit: Screening and diagnostic outcomes. *International journal of pediatric otorhinolaryngology*, *84*, 124-131.

De Sousa, K. C., Swanepoel, D. W., Moore, D. R., Myburgh, H. C., & Smits, C. (2020). Improving Sensitivity of the Digits-In-Noise Test Using Antiphasic Stimuli. *Ear and Hearing*, *41*(2), 442-450.

Denys, S., De Laat, J., Dreschler, W., Hofmann, M., van Wieringen, A., & Wouters, J. (2019). Language-Independent Hearing Screening Based on Masked Recognition of Ecological Sounds. *Trends in Hearing*, *23*, 2331216519866566.

Dimitriou, A., Perisanidis, C., Chalkiadakis, V., Marangoudakis, P., Tzagkaroulakis, A., & Nikolopoulos, T. P. (2016). The universal newborn hearing screening program in a public hospital: The importance of the day of examination. *International journal of pediatric otorhinolaryngology*, *91*, 90-93.

Downs, M. P., & Hemenway, W. G. (1969). Report on the Hearing Screening of 17.000 Neonates. *International Audiology*, *8*(1), 72-76.

EFCNI, Oudesluys-Murphy, A. M., van Wassenaer-Leemhuis, A., Wolke, D., & van Straaten, H. L. M. (2018). *European Standards of Care for Newborn Health: Hearing screening*. Retrieved December 5 from https://newborn-health-standards.org/hearing-screening/

Ewing, A. W. (1955). The sweep-frequency method of making screening tests of the hearing of schoolchildren. *Br Med J*, *1*(4904), 41-42.

Ewing, I. R., & Ewing, A. W. G. (1944). The Ascertainment of Deafness in Infancy and Early Childhood. *The Journal of Laryngology & Otology*, *59*(9), 309-333.

Fan, J. Y., Chen, L. S., Lai, J. C., Chen, M. K., & Chen, H. C. (2010). A pre-paid newborn hearing screening programme: a community-based study. *B-ENT*, *6*(4), 265-269.

Finitzo, T., & Grosse, S. (2003). Quality monitoring for early hearing detection and intervention programs to optimize performance [https://doi.org/10.1002/mrdd.10062]. *Mental Retardation and Developmental Disabilities Research Reviews*, *9*(2), 73-78.

Fisch, L. (1981). Development of School Screening Audiometry. *British Journal of Audiology*, *15*(2), 87-95.

Fortnum, H., & Davis, A. (1997). Epidemiology of Permanent Childhood Hearing Impairment in Trent Region, 1985–1993. *British Journal of Audiology*, *31*(6), 409-446.

Fortnum, H., Ukoumunne, O., Hyde, C., Taylor, R., Ozolins, M., Errington, S., . . . Moody, J. (2016). A programme of studies including assessment of diagnostic accuracy of school hearing screening tests and a cost-effectiveness model of school entry hearing screening programmes. *Health Technology Assessment*, *20*(36), 1-178.

Fortnum, H. M., Summerfield, A. Q., Marshall, D. H., Davis, A. C., & Bamford, J. M. (2001). Prevalence of permanent childhood hearing impairment in the United Kingdom and implications for universal neonatal hearing screening: questionnaire based ascertainment study. *Bmj*, *323*(7312), 536-540.

Friderichs, N., Swanepoel, D., & Hall, J. W. (2012). Efficacy of a community-based infant hearing screening program utilizing existing clinic personnel in Western Cape, South Africa. *International journal of pediatric otorhinolaryngology*, *76*(4), 552-559.

Gabbard, S. A., Northern, J. L., & Yoshinaga-Itano, C. (1999). Hearing screening in newborns under 24 hours of age. *Seminars in Hearing*, *20*(4), 291-305.

Gaffney, M., Eichwald, J., Grosse, S. D., & Mason, C. A. (2010). Identifying Infants with Hearing Loss --- United States, 1999--2007. *Weekly*, *59*(8), 220-223.

Gallus, R., Rizzo, D., De Luca, L. M., Melis, A., Kihlgren, C., Parente, P., . . . Conti, G. (2020). Does the involvement of first-year residents have a negative impact on the performance of a newborn hearing screening program? *International journal of pediatric otorhinolaryngology*, *138*, 110270.

Gorga, M. P., Norton, S. J., Sininger, Y. S., Cone-Wesson, B., Folsom, R. C., Vohr, B. R., . . . Stephen, T. (2000). Identification of Neonatal Hearing Impairment: Distortion Product Otoacoustic Emissions during the Perinatal Period. *Ear and Hearing*, *21*(5).

Gravel, J., Berg, A., Bradley, M., Cacace, A., Campbell, D., Dalzell, L., . . . Prieve, B. (2000). New York State universal newborn hearing screening demonstration project: effects of screening protocol on inpatient outcome measures. *Ear Hear*, *21*(2), 131-140.

Gumbie, M., Parkinson, B., Dillon, H., Bowman, R., Song, R., & Cutler, H. (2022). Cost-Effectiveness of Screening Preschool Children for Hearing Loss in Australia. *Ear and Hearing*, *43*(3).

Hall, J. W., Baer, J. E., Chase, P. A., & Schwaber, M. K. (1994). Clinical Application of Otoacoustic Emissions: What do we Know about Factors Influencing Measurement and Analysis? *Otolaryngology–Head and Neck Surgery*, *110*(1), 22-38.

Hearing screening: considerations for implementation. (2021). World Health Organization.

Hecox, K., & Galambos, R. (1974). Brain Stem Auditory Evoked Responses in Human Infants and Adults. *Archives of Otolaryngology, 99*(1), 30-33.

Hergils, L. (2007). Analysis of measurements from the first Swedish universal neonatal hearing screening program. *International Journal of Audiology, 46*(11), 680-685.

Hirsch, A., & Kankkunen, A. (1974). High Risk History in the Identification of Hearing Loss in Newborns. *Scandinavian Audiology, 3*(4), 177-182.

Hrncic, N., Hatibovic, H., Goga, A., & Hodzic, D. (2019). Does an early discharge of a newborn influence the success of the newborn hearing screening in developing countries? A hospital based study. *Medicinski Glasnik Ljekarske Komore Zenickodobojskog Kantona, 16*(2), 01.

Hunter, L. L., Meinzen-Derr, J., Wiley, S., Horvath, C. L., Kothari, R., & Wexelblatt, S. (2016). Influence of the WIC Program on Loss to Follow-up for Newborn Hearing Screening. *Pediatrics, 138*(1), 07.

Hussain, D. M., Gorga, M. P., Neely, S. T., Keefe, D. H., & Peters, J. (1998). Transient Evoked Otoacoustic Emissions in Patients with Normal Hearing and in Patients with Hearing Loss. *Ear and Hearing, 19*(6).

Hyde, M. (2016). Principles and Methods of Newborn Hearing Screening. In A. M. Tharpe & R. Seewald (Eds.), *Comprehensive Handbook of Pediatric Audiology, Second Edition* (pp. 309-445). Plural Publishing.

Jewett, D. L., Romano, M. N., & Williston, J. S. (1970). Human Auditory Evoked Potentials: Possible Brain Stem Components Detected on the Scalp. *Science, 167*(3924), 1517-1518.

Johnsen, N. J., Bagi, P., & Elberling, C. (1983). Evoked Acoustic Emissions from the Human Ear: III. Findings in Neonates. *Scandinavian Audiology, 12*(1), 17-24.

Johnson, A., & Ashurst, H. (1990). Screening for sensorineural deafness by health visitors. The Steering Committee, Oxford Region Child Development Project. *Archives of Disease in Childhood, 65*(8), 841.

Johnson, J. L., White, K. R., Widen, J. E., Gravel, J. S., James, M., Kennalley, T., . . . Holstrum, J. (2005). A Multicenter Evaluation of How Many Infants With Permanent Hearing Loss Pass a Two-Stage Otoacoustic Emissions/Automated Auditory Brainstem Response Newborn Hearing Screening Protocol. *Pediatrics, 116*(3), 663-672.

Joint Committee on Infant Hearing. (1982). Position Statement 1982. *Pediatrics, 70*(3), 496-497.

Joint Committee on Infant Hearing. (1991). 1990 position statement. *ASHA, 33*(suppl 5), 3-6.

Joint Committee on Infant Hearing. (1995). Joint Committee on Infant Hearing 1994 Position Statement. *Pediatrics, 95*(1), 152-156.

Joint Committee on Infant Hearing. (2000). Year 2000 Position Statement. *American Journal of Audiology, 9*(1), 9-29.

Joint Committee on Infant Hearing. (2007). Year 2007 position statement: Principles and guidelines for early hearing detection and intervention programs. *Pediatrics, 120*(4), 898-921.

Joint Committee on Infant Hearing. (2019). Year 2019 Position Statement: Principles and Guidelines for Early Hearing Detection and Intervention Programs. *The Journal of Early Hearing Detection and Intervention*, *4*(2), 1-44.

Junker, K. S., Barr, B., Maliniemi, S., & Wasz-höckert, O. (1978). BOEL Screening: A Program for the early detection of communicative disorders: Preliminary Reports from a Study on 1 000 Finnish Infants. *Audiology*, *17*(1), 51-61.

Kanji, A., Khoza-Shangase, K., & Moroe, N. (2018). Newborn hearing screening protocols and their outcomes: A systematic review. *International journal of pediatric otorhinolaryngology*, *115*, 104-109.

Kelly, A. F., Kelly, P. K., & Shah, M. (2021). Auditory Brainstem Response Pass Rates Correlate with Newborn Hour of Life and Delivery Mode. *Journal of Pediatrics*, *230*, 100-105.

Kemp, D. T. (2002). Otoacoustic emissions, their origin in cochlear function, and use. *British Medical Bulletin*, *63*(1), 223-241.

Kemp, D. T., & Ryan, S. (1991). Otoacoustic Emission Tests in Neonatal Screening Programmes. *Acta Oto-Laryngologica*, *111*(sup482), 73-84.

Kolski, C., Le Driant, B., Lorenzo, P., Vandromme, L., & Strunski, V. (2007). Early hearing screening: what is the best strategy? *International Journal of Pediatric Otorhinolaryngology*, *71*(7), 1055-1060.

Korres, S., Balatsouras, D., Ferekidis, E., Gkoritsa, E., Georgiou, A., & Nikolopoulos, T. (2003). The effect of different 'pass-fail' criteria on the results of a newborn hearing screening program. *Orl; Journal of Oto-Rhino-Laryngology & its Related Specialties*, *65*(5), 250-253.

Korres, S., Balatsouras, D. G., Vlachou, S., Kastanioudakis, I. G., Ziavra, N. V., & Ferekidis, E. (2005). Overcoming difficulties in implementing a universal newborn hearing screening program. *Turkish Journal of Pediatrics*, *47*(3), 203-212.

Korres, S., Nikolopoulos, T., Ferekidis, E., Gotzamanoglou, Z., Georgiou, A., & Balatsouras, D. G. (2003). Otoacoustic emissions in universal hearing screening: which day after birth should we examine the newborns? *Orl; Journal of Oto-Rhino-Laryngology & its Related Specialties*, *65*(4), 199-201.

Korver, A. M., van Zanten, G. A., Meuwese-Jongejeugd, A., van Straaten, H. L., & Oudesluys-Murphy, A. M. (2012). Auditory neuropathy in a low-risk population: a review of the literature. *International journal of pediatric otorhinolaryngology*, *76*(12), 1708-1711.

Korver, A. M. H., Konings, S., Dekker, F. W., Beers, M., Wever, C. C., Frijns, J. H. M., . . . Decibel Collaborative Study Group, f. t. (2010). Newborn Hearing Screening vs Later Hearing Screening and Developmental Outcomes in Children With Permanent Childhood Hearing Impairment. *JAMA*, *304*(15), 1701-1708.

Lemons, J., Fanaroff, A., Stewart, E. J., Bentkover, J. D., Murray, G., & Diefendorf, A. (2002). Newborn Hearing Screening: Costs of Establishing a Program. *Journal of Perinatology*, *22*(2), 120-124.

Levit, Y., Himmelfarb, M., & Dollberg, S. (2015). Sensitivity of the automated auditory brainstem response in neonatal hearing screening. *Pediatrics*, *136*(3), e641-e647.

Lin, H. C., Shu, M. T., Lee, K. S., Ho, G. M., Fu, T. Y., Bruna, S., & Lin, G. (2005). Comparison of hearing screening programs between one step with transient evoked otoacoustic emissions (TEOAE) and two steps with TEOAE and automated auditory brainstem response. *Laryngoscope*, *115*(11), 1957-1962.

Lin, H. C., Shu, M. T., Lee, K. S., Lin, H. Y., & Lin, G. (2007). reducing false positives in newborn hearing screening program: how and why. *Otology & Neurotology*, *28*(6), 788-792.

Lowell, E. L., Rushford, G., Hoversten, G., & Stoner, M. (1956). Evaluation of pure tone audiometry with preschool age children. *Journal of Speech and Hearing Disorders*, *21*(3), 292-302.

Lü, J., Huang, Z., Yang, T., Li, Y., Mei, L., Xiang, M., . . . Wu, H. (2011). Screening for delayed-onset hearing loss in preschool children who previously passed the newborn hearing screening. *International journal of pediatric otorhinolaryngology*, *75*(8), 1045-1049.

Marchal, B., van Belle, S., van Olmen, J., Hoerée, T., & Kegels, G. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, *18*(2), 192-212.

Mason, C., A., Gaffney, M., Green, D., R., & Grosse, S., D. (2008). Measures of Follow-Up in Early Hearing Detection and Intervention Programs: A Need for Standardization. *American Journal of Audiology*, *17*(1), 60-67.

Mauk, G. W., White, K. R., Mortensen, L. B., & Behrens, T. R. (1991). The Effectiveness of Screening Programs Based on High-Risk Characteristics in Early Identification of Hearing Impairment. *Ear and Hearing*, *12*(5).

Mazerolle, M. J. (2020). *AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c). R package version 2.3-1.* https://cran.r-project.org/package=AICcmodavg

Mehl, A. L., & Thomson, V. (1998). Newborn Hearing Screening: The Great Omission. *Pediatrics*, *101*(1), e4-e4.

Mehl, A. L., & Thomson, V. (2002). The Colorado newborn hearing screening project, 1992-1999: on the threshold of effective population-based universal newborn hearing screening. *Pediatrics*, *109*(1), E7.

Mincarone, P., Leo, C. G., Sabina, S., Costantini, D., Cozzolino, F., Wong, J. B., & Latini, G. (2015). Evaluating reporting and process quality of publications on UNHS: a systematic review of programmes. *BMC pediatrics*, *15*, 86-86.

Moeller, M. P. (2000). Early Intervention and Language Development in Children Who Are Deaf and Hard of Hearing. *Pediatrics*, *106*(3), e43-e43.

Moeller, M. P., Tomblin, J. B., Yoshinaga-Itano, C., Connor, C. M., & Jerger, S. (2007). Current State of Knowledge: Language and Literacy of Children with Hearing Impairment. *Ear and Hearing*, *28*(6).

Mott, A., & Emond, A. (1994). What is the role of the distraction test of hearing? *Archives of Disease in Childhood*, *70*(1), 10.

Neumann, K., Chadha, S., Tavartkiladze, G., Bu, X., & White, K. R. (2019). Newborn and Infant Hearing Screening Facing Globally Growing Numbers of People Suffering from Disabling Hearing Loss. *International Journal of Neonatal Screening*, *5*(1), 7.

Norton, S. J., Gorga, M. P., Widen, J. E., Folsom, R. C., Sininger, Y., Cone-Wesson, B., . . . Fletcher, K. (2000). Identification of Neonatal Hearing Impairment: Evaluation of Transient Evoked Otoacoustic Emission, Distortion Product Otoacoustic Emission, and Auditory Brain Stem Response Test Performance. *Ear and Hearing*, *21*(5), 508-528.

Norton, S. J., Gorga, M. P., Widen, J. E., Vohr, B. R., Folsom, R. C., Sininger, Y. S., . . . Fletcher, K. A. (2000). Identification of Neonatal Hearing Impairment: Transient Evoked Otoacoustic Emissions during the Perinatal Period. *Ear and Hearing*, *21*(5).

Olusanya, B. O. (2012). Neonatal hearing screening and intervention in resource-limited settings: an overview. *Arch Dis Child*, *97*(7), 654-659.

Olusanya, B. O. (2015). Screening for neonatal deafness in resource-poor countries: challenges and solutions. *Research and Reports in Neonatology*, *5*, 51-64.

Ong, K. M. C., Rivera, A. S., Chan, A. L., & Chiong, C. M. (2020). Determining concordance and cost impact of otoacoustic emission and automated auditory brainstem response in newborn hearing screening in a tertiary hospital. *International journal of pediatric otorhinolaryngology*, *128*, 109704.

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., . . . McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *Bmj*, *372*, n160.

Park, S. K., Chang, J., Chung, Y. S., & Oh, S. H. (2020). Analysis of the effectiveness of coupon-mediated newborn hearing screening program through comparison of two government-funded pilot projects in South Korea. *International journal of pediatric otorhinolaryngology*, *136*, 110256.

Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. Sage.

Pimperton, H., Blythe, H., Kreppner, J., Mahon, M., Peacock, J. L., Stevenson, J., . . . Kennedy, C. R. (2014). The impact of universal newborn hearing screening on long-term literacy outcomes: a prospective cohort study. *Archives of Disease in Childhood*, archdischild-2014-307516.

Pirozzo, S., Papinczak, T., & Glasziou, P. (2003). Whispered voice test for screening for hearing impairment in adults and children: systematic review. *Bmj*, *327*(7421), 967.

Pitathawatchai, P., Chaichulee, S., Wannaro, W., & Pongprawat, P. (2022). Cost-effectiveness analysis on implementing newborn hearing screening programmes in a low- to middle-income country. *International Journal of Audiology*, 1-10.

Potgieter, J.-M., Swanepoel, D. W., Myburgh, H. C., Hopper, T. C., & Smits, C. (2016). Development and validation of a smartphone-based digits-in-noise hearing test in South African English. *International Journal of Audiology*, *55*(7), 405-411.

Prescott, C. A. J., Omoding, S. S., Fermor, J., & Ogilvy, D. (1999). An evaluation of the 'voice test' as a method for assessing hearing in children with particular reference to the situation in developing countries. *International journal of pediatric otorhinolaryngology*, *51*(3), 165-170.

Prieve, B., & Fitzgerald, T. (2014). Otoacoustic emissions. In J. Katz, L. Medwetsky, R. Burkard, & L. J. Hood (Eds.), *Handbook of clinical audiology* (pp. 357-379). Taylor and Francis Inc.

Prieve, B. A., Gorga, M. P., Schmidt, A., Neely, S., Peters, J., Schultes, L., & Jesteadt, W. (1993). Analysis of transient-evoked otoacoustic emissions in normal-hearing and hearing-impaired ears. *The Journal of the Acoustical Society of America*, *93*(6), 3308-3319.

Prince, C. B., Miyashiro, L., Weirather, Y., & Heu, P. (2003). Epidemiology of early hearing loss detection in Hawaii. *Pediatrics*, *111*(5 Pt 2), 1202-1206.

R Core Team. (2021). *R: A language and environment for statistical computing.* In R Foundation for Statistical Computing. https://www.R-project.org/

Rance, G., Beer, D. E., Cone-Wesson, B., Shepherd, R. K., Dowell, R. C., King, A. M., . . . Clark, G. M. (1999). Clinical Findings for a Group of Infants and Young Children with Auditory Neuropathy. *Ear and Hearing*, *20*(3).

Rance, G., & Starr, A. (2017). Auditory neuropathy spectrum disorder. *Comprehensive handbook of pediatric audiology*, 227-246.

Ravi, R., Gunjawate, D. R., Yerraguntla, K., Lewis, L. E., Driscoll, C., & Rajashekhar, B. (2016). Follow-up in newborn hearing screening – A systematic review. *International journal of pediatric otorhinolaryngology*, *90*, 29-36.

Razak, A., Fard, D., Hubbell, R., Cohen, M., Hartman-Joshi, K., & Levi, J. R. (2021). Loss to Follow-Up After Newborn Hearing Screening: Analysis of Risk Factors at a Massachusetts Urban Safety-Net Hospital. *Ear and Hearing*, *42*(1).

Reznik, R., Starte, D., & Morey, S. U. E. (1985). Health screening at school entry—what is achieved? [https://doi.org/10.1111/j.1440-1754.1985.tb02124.x]. *Journal of Paediatrics and Child Health*, *21*(3), 159-162.

Russ, S. A., Hanna, D., DesGeorges, J., & Forsman, I. (2010). Improving Follow-up to Newborn Hearing Screening: A Learning-Collaborative Experience. *Pediatrics*, *126*(Supplement_1), S59-S69.

Scheepers, L. J., Swanepoel de, W., & Roux, T. (2014). Why parents refuse newborn hearing screening and default on follow-up rescreening--a South African perspective. *International Journal of Pediatric Otorhinolaryngology*, *78*(4), 652-658.

Sekhar, D. L., Zalewski, T. R., & Paul, I. M. (2013). Variability of State School-Based Hearing Screening Protocols in The United States. *Journal of Community Health*, *38*(3), 569-574.

Shang, Y., Hao, W., Gao, Z., Xu, C., Ru, Y., & Ni, D. (2016). An effective compromise between cost and referral rate: A sequential hearing screening protocol using TEOAEs and AABRs for healthy newborns. *International journal of pediatric otorhinolaryngology*, *91*, 141-145.

Sharma, R., Gu, Y., Ching, T. Y. C., Marnane, V., & Parkinson, B. (2019). Economic Evaluations of Childhood Hearing Loss Screening Programmes: A Systematic Review and Critique. *Applied Health Economics and Health Policy*, *17*(3), 331-357.

Sharma, R., Gu, Y., Sinha, K., Ching, T. Y. C., Marnane, V., Gold, L., . . . Parkinson, B. (2022). An Economic Evaluation of Australia's Newborn Hearing Screening Program: A Within-Study Cost-Effectiveness Analysis. *Ear and Hearing*, *43*(3).

Shoup, A. G., Owen, K. E., Jackson, G., & Laptook, A. (2005). The Parkland Memorial Hospital experience in ensuring compliance with Universal Newborn Hearing Screening follow-up. *Journal of Pediatrics*, *146*(1), 66-72.

Simmons, F. B., & Russ, F. N. (1974). Automated Newborn Hearing Screening, the Crib-o-gram. *Archives of Otolaryngology*, *100*(1), 1-7.

Sininger, Y. S. (1993). Auditory brain stem response for objective measures of hearing. *Ear Hear*, *14*(1), 23-30.

Sininger, Y. S., Abdala, C., & Cone-Wesson, B. (1997). Auditory threshold sensitivity of the human neonate as measured by the auditory brainstem response. *Hearing Research*, *104*(1), 27-38.

Skarżyński, H., & Piotrowska, A. (2012). Screening for pre-school and school-age hearing problems: European Consensus Statement. *International journal of pediatric otorhinolaryngology*, *76*(1), 120-121.

Skoloudik, L., Mejzlik, J., Janouch, M., Drsata, J., Vodicka, J., & Chrobok, V. (2020). Hearing screenings for preschool children: A comparison between whispered voice and pure tone audiogram tests. *International journal of pediatric otorhinolaryngology*, *130*, 109798.

Sloot, F., Hoeve, H. L. J., de Kroon, M. L. A., Goedegebure, A., Carlton, J., Griffiths, H. J., & Simonsz, H. J. (2015). Inventory of current EU paediatric vision and hearing screening programmes. *Journal of Medical Screening*, *22*(2), 55-64.

Smits, C., Theo Goverts, S., & Festen, J. M. (2013). The digits-in-noise test: Assessing auditory speech recognition abilities in noise. *The Journal of the Acoustical Society of America*, *133*(3), 1693-1706.

Starr, A., Picton, T. W., Sininger, Y., Hood, L. J., & Berlin, C. I. (1996). Auditory neuropathy. *Brain*, *119*(3), 741-753.

Stenfeldt, K. (2018). Preschool hearing screening in Sweden. An evaluation of current practices and a presentation of new national guidelines. *International journal of pediatric otorhinolaryngology*, *110*, 70-75.

Stewart, D. L., Mehl, A., Hall, J. W., 3rd, Thomson, V., Carroll, M., & Hamlett, J. (2000). Universal newborn hearing screening with automated auditory brainstem response: a multisite investigation. *Journal of perinatology : official journal of the California Perinatal Association*, *20*(8 Pt 2), S128-131.

Subbiah, K., Mason, C. A., Gaffney, M., & Grosse, S. D. (2018). Progress in Documented Early Identification and Intervention for Deaf and Hard of Hearing Infants: CDC's Hearing Screening and Follow-up Survey, United States, 2006-2016. *Journal of early hearing detection and intervention*, *3*(2), 1-7.

Tabrizi, A. G., Asadi, M., Barati, B., & Rabori, M. S. (2017). BIRTH BY CESAREAN DELIVERY ON NEWBORN HEARING SCREENING TEST: A RETROSPECTIVE STUDY. *International Journal of Life Science and Pharma Research*, *7*(4), L26-L29.

Therneau, T. (2022). *_A Package for Survival Analysis in R_. R package version 3.3-0.* https://CRAN.R-project.org/package=survival

Thomson, V., & Yoshinaga-Itano, C. (2018). The Role of Audiologists in Assuring Follow-Up to Outpatient Screening in Early Hearing Detection and Intervention Systems. *American Journal of Audiology*, *27*(3), 283-293.

Tobe, R. G., Mori, R., Huang, L., Xu, L., Han, D., & Shibuya, K. (2013). Cost-Effectiveness Analysis of a National Neonatal Hearing Screening Program in China: Conditions for the Scale-Up. *PloS one*, *8*(1), e51990.

Tomblin, J. B., Oleson, J., Ambrose, S. E., Walker, E. A., & Moeller, M. P. (2020). Early Literacy Predictors and Second-Grade Outcomes in Children Who Are Hard of Hearing. *Child development*, *91*(1), e179-e197.

Uhlén, I. M., Mackey, A. R., & Rosenhall, U. (2020). Prevalence of childhood hearing impairment in the County of Stockholm – a 40-year perspective from Sweden and other high income countries. *International Journal of Audiology*, *59*(11), 866-873.

Uilenburg, N., Kauffman-de Boer, M., van der Ploeg, K., Oudesluys-Murphy, A. M., & Verkerk, P. (2009). An implementation study of neonatal hearing screening in the Netherlands. *International Journal of Audiology*, *48*(3), 108-116.

Uus, K., & Bamford, J. (2006). Effectiveness of Population-Based Newborn Hearing Screening in England: Ages of Interventions and Profile of Cases. *Pediatrics*, *117*(5), e887-e893.

Uus, K., Bamford, J., & Taylor, R. (2006). An analysis of the costs of implementing the National Newborn Hearing Screening Programme in England. *Journal of Medical Screening*, *13*(1), 14-19.

Verkleij, M. L., Heijnsdijk, E. A. M., Bussé, A. M. L., Carr, G., Goedegebure, A., Mackey, A. R., . . . de Koning, H. J. (2021). Cost-Effectiveness of Neonatal Hearing Screening Programs: A Micro-Simulation Modeling Analysis. *Ear Hear*, *42*(4), 909-916.

Vernier, L. S., Schneider, K. L., Zanini, C., Paniz, T., & Levandowski, D. C. (2021). Delivery Route and the Outcome of Newborn Hearing Screening of Full-Term Neonates Born in a Public Maternal-Infant Hospital in the South of Brazil [Article; Early Access]. *International Archives of Otorhinolaryngology*.

Vernon, M. (1969). Sociological and Psychological Factors Associated with Hearing Loss. *Journal of Speech and Hearing Research*, *12*(3), 541-563.

Vohr, B. R., Moore, P. E., & Tucker, R. J. (2002). Impact of Family Health Insurance and Other Environmental Factors on Universal Hearing Screen Program Effectiveness. *Journal of Perinatology*, *22*(5), 380-385.

Vohr, B. R., Oh, W., Stewart, E. J., Bentkover, J. D., Gabbard, S., Lemons, J., . . . Pye, R. (2001). Comparison of costs and referral rates of 3 universal newborn hearing screening protocols. *J Pediatr*, *139*(2), 238-244.

Vohr, B. R., White, K. R., Maxon, A. B., & Johnson, M. J. (1993). Factors affecting the interpretation of transient evoked otoacoustic emission results in neonatal hearing screening. *Seminars in Hearing*, *14*(1), 57-72.

Wake, M., Hughes, E. K., Collins, C. M., & Poulakis, Z. (2004). Parent-Reported Health-Related Quality of Life in Children With Congenital Hearing Loss: A Population Study. *Ambulatory Pediatrics*, *4*(5), 411-417.

Walker, E. A., Holte, L., McCreery, R. W., Spratford, M. h., Page, T., & Moeller, M. P. (2015). The Influence of Hearing Aid Use on Outcomes of Children With Mild Hearing Loss. *Journal of Speech, Language, and Hearing Research*, *58*(5), 1611-1625.

Walker, E. A., Sapp, C., Dallapiazza, M., Spratford, M., McCreery, R. W., & Oleson, J. J. (2020). Language and Reading Outcomes in Fourth-Grade Children With Mild Hearing Loss Compared to Age-Matched Hearing Peers. *Language, Speech, and Hearing Services in Schools*, *51*(1), 17-28.

Walker, E. A., Ward, C., Oleson, J. J., Sapp, C., McCreery, R., Tomblin, J. B., & Moeller, M. P. (2022). Language growth in children with mild to severe hearing loss who received early intervention by 3 months or 6 months of age. *Journal of early hearing detection and intervention*, *7*(1), 1-10.

Watkin, P., & Baldwin, M. (2012). The longitudinal follow up of a universal neonatal hearing screen: The implications for confirming deafness in childhood. *International Journal of Audiology*, *51*(7), 519-528.

Watkin, P. M., & Baldwin, M. (1999). Confirmation of deafness in infancy. *Archives of Disease in Childhood*, *81*(5), 380.

Watson, D. R., McClelland, R. J., & Adams, D. A. (1996). Auditory brainstem response screening for hearing loss in high risk neonates. *International journal of pediatric otorhinolaryngology*, *36*(2), 147-183.

Weichbold, V., Nekahm-Heis, D., & Welzl-Mueller, K. (2006). Universal Newborn Hearing Screening and Postnatal Hearing Loss. *Pediatrics*, *117*(4), e631-e636.

Wells, G. A., Shea, B., O'Connell, D. a., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2000). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. In: Oxford.

Wessex Universal Neonatal Hearing Screening Trial Group. (1998). Controlled trial of universal neonatal screening for early identification of permanent childhood hearing impairment. *The Lancet*, *352*(9145), 1957-1964.

Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., . . . Bossuyt, P. M. (2011). QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of internal medicine*, *155*(8), 529-536.

Wilson, B. S., Tucci, D. L., Merson, M. H., & O'Donoghue, G. M. (2017). Global hearing health care: new findings and perspectives. *Lancet*, *390*(10111), 2503-2515.

Wilson, J. M. G., & Jungner, G. (1968). *Principles and practice of screening for disease*. World Health Organization.

Wong, C. L., Ching, T. Y. C., Cupples, L., Button, L., Leigh, G., Marnane, V., . . . Martin, L. (2017). Psychosocial Development in 5-Year-Old Children With Hearing Loss Using Hearing Aids or Cochlear Implants. *Trends in Hearing*, *21*, 2331216517710373.

Wood, S., Davis, A. C., & McCormick, B. (1997). Changing Performance of the Health Visitor Distraction Test When Targeted Neonatal Screening is Introduced into a Health District. *British Journal of Audiology*, *31*(1), 55-61.

Wood, S. A., Sutton, G. J., & Davis, A. C. (2015). Performance and characteristics of the Newborn Hearing Screening Programme in England: The first seven years. *Int J Audiol*, *54*(6), 353-358.

World Health Organization. (2016). *Childhood hearing loss: Strategies for prevention and care*. https://apps.who.int/iris/handle/10665/204632

World Health Organization. (2021). *World report on hearing*. World Health Organization. License: CC BY-NC-SA 3.0 IGO.

Yong, M., Liang, J., Ballreich, J., Lea, J., Westerberg, B. D., & Emmett, S. D. (2020). Cost-effectiveness of School Hearing Screening Programs: A Scoping Review. *Otolaryngology–Head and Neck Surgery*, *162*(6), 826-838.

Yong, M., Panth, N., McMahon, C. M., Thorne, P. R., & Emmett, S. D. (2020). How the World's Children Hear: A Narrative Review of School Hearing Screening Programs Globally. *OTO Open*, *4*(2), 2473974X20923580.

Yoshinaga-Itano, C. (2003). Early intervention after universal neonatal hearing screening: Impact on outcomes [https://doi.org/10.1002/mrdd.10088]. *Mental Retardation and Developmental Disabilities Research Reviews*, *9*(4), 252-266.

Yoshinaga-Itano, C. (2004). Levels of evidence: universal newborn hearing screening (UNHS) and early hearing detection and intervention systems (EHDI). *Journal of Communication Disorders*, *37*(5), 451-465.

Yoshinaga-Itano, C., Sedey, A. L., Coulter, D. K., & Mehl, A. L. (1998). Language of Early- and Later-identified Children With Hearing Loss. *Pediatrics*, *102*(5), 1161-1171.

Zeitlin, W., Auerbach, C., Mason, S. E., Spivak, L. G., & Reiter, B. (2017). Factors Related to Not Following Up with Recommended Testing in the Diagnosis of Newborn Hearing Loss. *Health & Social Work*, *42*(1), 24-31.

Zeitlin, W., McInerney, M., Aveni, K., Scheperle, R., & DeCristofano, A. (2021). Maternal Factors Predicting Loss to Follow-Up from Newborn Hearing Screenings in New Jersey. *Health & Social Work*, *46*(2), 115-124.

10 APPENDIX

Table A-1. Test method used for each step for low-risk and high-risk infants across countries or regions surveyed with the EUSCREEN questionnaire.

	Low-risk protocol	High-risk protocol
OAE	Russia*	Romania Russia†
OAE, OAE	Romania* Bosnia & Herzegovina (Tuzla Canton) Latvia Lithuania Serbia (Belgrade) Moldova (Chisinau) Belgium (Wallonia-Brussels Federation) Austria (Upper Austria) Croatia Czechia (East Bohemia)* Poland Luxembourg Spain (Aut. Comm. Valencia)* Bulgaria* Portugal (Lisbon) Slovakia Slovenia Switzerland	Bosnia & Herzegovina (Tuzla Canton) Latvia Lithuania Malta North Macedonia (Skopje) Moldova (Chisinau) Poland† Serbia
OAE, OAE, OAE	Spain (Princp. Asturias)	
OAE, aABR	China Germany (Westphalia-Lippe)**	
OAE, OAE, aABR	Faroe Islands Estonia* Cyprus (Southern part) Israel* Greece Iceland Netherlands Italy (Veneto Region) Finland Ireland England (SE London) Turkey* Denmark** France (Ile de France)**	Faroe Islands Estonia
OAE, OAE, OAE, aABR	Sweden (Stockholm Region)	
aABR	Hungary*	Spain (Aut. Comm. Valencia) Portugal Slovenia Switzerland Hungary

		China
aABR, aABR	Belgium (Flanders)	Austria (Upper Austria) Croatia Czechia (East Bohemia) Luxembourg Belgium (Flanders) Germany (Westphalia-Lippe) Greece Iceland
aABR, aABR, aABR		France (Ile de France)
OAE+aABR		Bulgaria Slovakia Spain (Princp. Asturias) Cyprus (Southern part) Israel Denmark Sweden (Stockholm Region) England (SE London) Turkey
OAE+aABR, OAE+aABR		Italy (Veneto)
Varies	India	Finland India
Direct referral to diagnostic assessment		Belgium (Wallonia-Brussels Federation)
No screening	Albania Kosovo Malawi Rwanda Montenegro Malta North Macedonia (Skopje Region)	Albania Kosovo Malawi Rwanda Montenegro

*Protocol allows for multiple screens within one step; other regions to not specify whether multiple screens are performed or not; **aABR may be used in place of OAE in some hospitals; †Referral is made to diagnostic assessment regardless of screening result