JEHDI
The Journal of Early Hearing Detection and Intervention
NCHAM National Center for Hearing Assessment and Management Utah State University™
2021; 6(1): 12-31

Developing a Spoken Language Outcome Monitoring Procedure for a Canadian Early Hearing Detection and Intervention Program: Process and Recommendations

Olivia Daub, MSc[1]
Janis Oram Cardy, PhD[2]
[1]Graduate Program in Health and Rehabilitation Sciences, The University of Western Ontario, London, Ontario, Canada
[2]School of Communication Sciences and Disorders and the National Centre for Audiology, The University of Western Ontario, London, Ontario, Canada

Abstract

Purpose: Routine spoken language outcome monitoring is one component of Early Hearing Detection and Intervention (EHDI) programs for children who are hard of hearing and learning a spoken language. However, there is no peer-reviewed research that documents how spoken language outcome monitoring may be achieved, or what processes EHDI programs can use to develop these procedures. The present article describes the process used by a Canadian EHDI program and the final recommendations that were developed from this process.

Methodology: Through consultation with the program's stakeholders, consideration of the Joint Committee on Infant Hearing's recommendations, and drawing on our own expertise in spoken language assessment, we developed an overall framework for monitoring spoken language. Based on the needs of the EHDI program, we conducted a scoping review and critical appraisal of norm-referenced tests to identify candidate tests to use within this framework.

Results: We recommended a two-pronged assessment approach to measuring spoken language outcomes, including program-level assessment and individual vulnerability testing. We identified several tests that have been previously used to measure spoken language outcomes. There was little consistency in how tests were used across studies with no clear indicators as to which tests are the most appropriate to accomplish for which outcome monitoring purposes.

Conclusions: This article reports on the framework and tests used by a Canadian EHDI program to accomplish spoken language outcome monitoring. We highlight different factors that need to be considered when designing spoken language outcome monitoring procedures and the complexity in doing so. Future work evaluating the effectiveness and feasibility of our recommendations is warranted.

Keywords: Spoken language outcome monitoring; Program Evaluation

Acronyms: CASL = Comprehensive Assessment of Spoken Language; CDI = Child Development Inventory; CELF = Comprehensive Evaluation of Language Fundamentals; COSMIN = Consensus Based Standards for the Selection of Health Status Measurement Instruments; DEAP = Diagnostic Evaluation of Articulation and Phonology; EHDI = Early Hearing Detection and Intervention; EOWPVT = Expressive One Word Vocabulary Test; EVT = Expressive Vocabulary Test; GFTA = Goldman-Fristoe Test of Articulation; IHP = Infant Hearing Program; KLPA = Khan-Lewis Phonological Analysis; MBCDI = MacArthur Bates Communicative Development Inventories; (M)CDI = (Minnesota) Child Development Inventory; MSEL = Mullen Scales of Early Learning; PLAI = Preschool Language Assessment Inventory; PLS = Preschool Language Scale; PPVT= Peabody Picture Vocabulary Test; SLP = speech language pathologist; TACL = Test of Auditory Comprehension of Language, VABS = Vineland Adaptive Behavior Scales

Acknowledgements: The authors have no conflicts of interest to declare. This work was funded by the Ontario Ministry of Children, Community and Social Services.

The authors would like to thank the speech-language pathologists, audiologists, and program managers who contributed to the development of these procedures and recommendations. We would also like to thank Kelsi Breton for her work in evaluating articles for inclusion and exclusion.

Correspondence concerning this article should be addressed to: Olivia Daub, Graduate Program in Health and Rehabilitation Sciences, The University of Western Ontario, Elborn College, London, Ontario, Canada, N6G 1H1. Email: odaub@uwo.ca

Early Hearing Detection and Intervention (EHDI) programs provide family centered support in the pursuit of typical language development (whether signed or spoken) for children who are deaf and hard-of-hearing (DHH; Moeller et al., 2013). For families who elect to pursue language in a spoken modality, EHDI programs have been demonstrated to improve spoken language outcomes (Ching, Day et al., 2013; Moeller, 2000;

Tomblin et al., 2015). Recent research has identified that interventions provided through EHDI programs such as early amplification, high levels of audibility, and support for consistent hearing aid use, are significant predictors of eventual spoken language outcomes and growth in spoken language over time (Tomblin et al., 2015). Comprehensive EHDI programs are gaining increasing international support, and international recommendations have been developed to guide their implementation (Moeller et al., 2013; Joint Committee on Infant Hearing, 2013, 2019). The Joint Committee on Infant Hearing (JCIH) has worked for many years to establish guidelines to ensure consistent and equitable service for children who are DHH and enrolled in different EHDI programs across the United States, and their work has set a standard for EHDI programs worldwide (e.g., the Canadian Infant Hearing Task Force endorses these recommendations). One of the committee's activities has been the publication of position statements summarizing the current state of the evidence in infant hearing and providing preferred practice recommendations on early identification and intervention for children who are DHH.

Of interest to the present article are JCIH recommendations for routine outcome monitoring of children enrolled in EHDI programs, specifically the monitoring of language outcomes. Because a central aim of EHDI programs is to prevent developmental delays associated with permanent childhood hearing loss, the recommendation for routine monitoring of spoken language development (when this is the mode of communication chosen by the family) is intended to ensure that "a child's developmental progress is comparable with his or her hearing peers" (JCIH, 2007, p. 909) and within 1 *SD* of their age or cognitive development on norm-referenced spoken language testing (JCIH, 2013). To meet this expectation, the JCIH recommends that policymakers, service providers, and family members use the results of routine spoken language outcome monitoring to support decision making. For instance, results from spoken language monitoring should be used to inform program evaluation and quality assurance at the program level, support comparison between EHDI programs using national databases, inform intervention planning at the level of the individual child and family, and determine whether a child is or is not meeting developmental milestones (JCIH, 2013, 2019).

However, there is no clear guidance on how EHDI programs ought to accomplish spoken language outcome monitoring, and the concept of spoken language outcome monitoring is poorly defined. Spoken language encompasses a wide range of inter-related skills, some of which a child may or may not struggle with at different ages. Nor do recommendations connect assessment purposes with tests or propose solutions to overcome the psychometric challenges associated with defining acceptable outcomes. Identifying the intended purpose(s) of conducting routine measurement of spoken language outcomes is an essential consideration in selecting the assessment approaches and which tests to use (Daub et al., in press), because different tests may be better suited to different purposes. Furthermore, not all tests are validated to support multiple decisions (Daub et al., 2019) and some assessment purposes are at psychometric odds with one another. For instance, the appropriate composition of a normative sample changes if the test is being used for absolute purposes (i.e., determining whether a child is below age expectations) or relative purposes (determining the severity of a spoken language disorder; Peña et al., 2006). As outlined by JCIH (2007, p. 909), "the primary purpose of regular developmental monitoring is to provide valuable information to parents about the rate of their child's development as well as programmatic feedback concerning curriculum decisions." These two decisions (i.e., information about rate of development and programmatic feedback) imply two conflicting purposes: measurement that is sensitive to an individual child's growth over time and measurement that is comparable between all children in a program. In speech-language pathology, it is traditionally advised to avoid measuring growth with norm-referenced tests because these tests are inherently broad, robust, and stable measures of spoken language constructs that aren't designed to be sensitive to change in language ability (McCauley & Swisher, 1984). However, relatively new statistics (e.g., item response theory derived scores such as growth scale values) that can be used to measure change over time are increasingly being reported in norm-referenced tests, although these are not yet commonplace (Daub et al., 2017; Daub et al., 2019). Comparing results between groups of children for the purpose of evaluating the broader EHDI program, however, requires that all children in the program are assessed at regular intervals with a consistent measure so that norm-referenced results can be compared.

The present project was born out of our efforts to support a Canadian EHDI program, the Ontario Infant Hearing Program (IHP), which serves children from birth to age 6, in developing a spoken language outcome monitoring procedure that would allow them to fulfill best-practice recommendations. The IHP was developed in 2002 and is a publicly funded EHDI program. The IHP provides universal newborn hearing screening services to all babies born in Ontario and intervention services to children with permanent hearing loss up to the age of 6 years. Spoken language development services for children in the IHP are provided by the publicly funded Ontario Preschool Speech and Language Program until they transition to school services, which can start as early as 3 or 4 years for those who attend junior kindergarten, but does not occur until 6 years of age for others. The IHP provides language development support in the primary language modality (either signed or spoken) as chosen by the family (Moeller et al., 2013) and may include technological intervention (e.g., hearing aids), sign language consultation, or spoken language intervention through speech-language pathology services. However, it is not the case that families are committed to selecting one language modality. Rather, given the publicly funded nature of the program, the

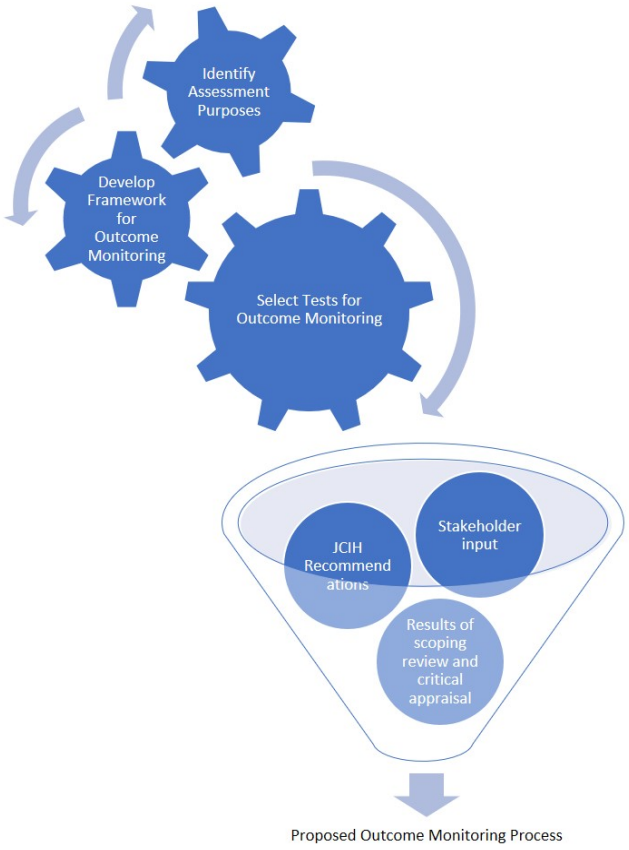
IHP provides funding for families to access services to support a primary language modality and families may pursue additional, privately funded services if, for instance, they wish to raise their child in a bimodal bilingual environment. Similarly, children in the IHP who are learning spoken language may also be raised in homes with two spoken languages. In cases where cochlear implantation is indicated, families access support through a collaboration with a separate publicly funded program and may not be followed by the IHP specifically. As a result, the present article focuses specifically on children who are hard of hearing (HH) and not children who are candidates for cochlear implantation. The IHP aligns its expectations closely with the recommendations put forth by the Canadian Infant Hearing Task Force and the JCIH. Currently, Ontario is one of six Canadian provinces/territories judged to be sufficiently meeting EHDI program standards (Canadian Infant Hearing Task Force, 2019).

Since 2009, spoken language outcome monitoring in the IHP has been conducted using the *Preschool Language Scale, 4th ed* (PLS-4; Zimmerman et al., 2006) every 6 months (JCIH 2007; 2013). Outcomes were to be tracked for all children for whom families selected spoken language as a primary language modality. This group can include children learning spoken language only or in conjunction with a signed language. Our research team was previously contracted by the IHP to evaluate outcomes using PLS-4 data from two birth cohorts in the program (Daub, 2016; Daub et al., 2017) and were therefore familiar with the previous process, as well as elements of data collection and reporting that were inconsistently implemented across the program. For example, less than 50% of the children in the birth cohorts analyzed did not have PLS-4 scores in the database, and PLS-4 scores were inconsistently scored across children (Daub, 2016). Because the nature of our involvement with the PLS-4 data was post-hoc, it was unclear whether data collection issues stemmed from issues with administration of the PLS-4, data entry/management errors, or errors in extraction from the data management system. The amount of data that were missing for undocumented reasons highlighted the importance of improving upon the previous procedure to support program evaluation. Around the same time that our team was involved in evaluating the outcome data from previous cohorts, the PLS-4 fell out of print in favor of the *Preschool Language Scale, 5th ed* (PLS-5; Zimmerman et al., 2011). As a result, the IHP sought to confirm that the PLS-5 would be an adequate replacement, and to evaluate and reconsider their procedure if necessary. At the same time, speech-language pathologists (SLPs) raised concerns about the appropriateness of the PLS-4/PLS-5 and questioned the rationale for its selection.

This article reports on a series of program evaluation and quality improvement projects we conducted to facilitate the IHP's decision-making about a new spoken language outcome monitoring procedure. These projects began in 2014, and our initial recommendations were shared with the IHP in 2017. We begin by orienting the reader to the overall process we used to develop the procedure (see Figure 1). This includes identifying the IHP's assessment purposes, developing a framework for assessing outcomes, and identifying tests to use in the framework. We then report on how we identified tests that appropriately fit within the framework, while also balancing needs at the level of both the program and the individual service providers and families.

Figure 1

Process for Developing Proposed Outcome Monitoring Process



Step 1: Identifying Assessment Purposes

The IHP's Assessment Purposes at the Program Level

Our main priority was to collect and maintain data within a provincial database that was appropriate for (a) evaluating the overall expressive and receptive spoken language outcomes of children in the IHP as a group to demonstrate the effectiveness of the IHP, (b) modeling children's spoken language growth over time to identify ages/stages of development where additional support might be needed, (c) identifying predictors of better, or worse, spoken language outcomes to support quality improvement initiatives, and (d) identifying whether there are differences in outcomes across regions of the province to support resource allocation. IHP management was also cognizant of the importance of clinician's assessment purposes and minimizing the time and financial burden of spoken language outcome monitoring on service providers to the greatest extent possible. They were also interested in a procedure that could provide clinically useful data about individual children in addition to program-level evaluation.

The IHP's Assessment Purposes for Individual Children and Families

At the level of the individual child and family, routine assessment of speech and language development should (a) identify children who are performing below age expectations and thus require speech-language development services, (b) allow profiling areas of relative strength and weakness in individual children, thus enabling clinicians to set goals and tailor interventions to meet individual needs at different stages of the child's development, and (c) allow for evaluation of school readiness and anticipation of academic supports needed to ensure success upon school entry. Because children with permanent hearing loss have ongoing inconsistent access to auditory information, they are at greater risk for difficulties in certain areas of spoken language than others (Moeller, Tomblin, et al., 2007), even if they perform within age expectations on omnibus spoken language tests. Therefore, developing a procedure that is informative to intervention planning for individual children required an approach that probed more deeply than overall spoken language outcomes, specifically those domains of language that are (a) known to be at particular risk in children with permanent hearing loss and (b) predictive of future spoken language outcomes. For children with moderate to severe hearing loss, who are served by the IHP, there are certainly gaps in knowledge about development of specific spoken language domains (Moeller, Tomblin, et al., 2007), but some of the most vulnerable domains in children from birth to 6 years appear to be related to inconsistencies in auditory access, including:

1. Vocal development and canonical babbling in infancy (Moeller, Hoover, Putnam, Arbataitis, Bohnenkamp, Peterson, Wood, et al., 2007; Moeller, Hoover, Putnam, Arbataitis, Bohnenkamp, Peterson, Wood, Lewis, et al., 2007; Moeller, Tomblin, et al., 2007)
2. Syllable structure and early vocabulary in the toddler period (Moeller, Hoover, Putnam, Arbataitis, Bohnenkamp, Peterson, Wood, Lewis, et al., 2007)
3. Morphosyntactic difficulty, which is suspected to stem from underlying concerns with articulation and phonology (Moeller, Tomblin, et al., 2007)
4. Phonological awareness in the preschool/kindergarten period (Moeller, Tomblin, et al., 2007)

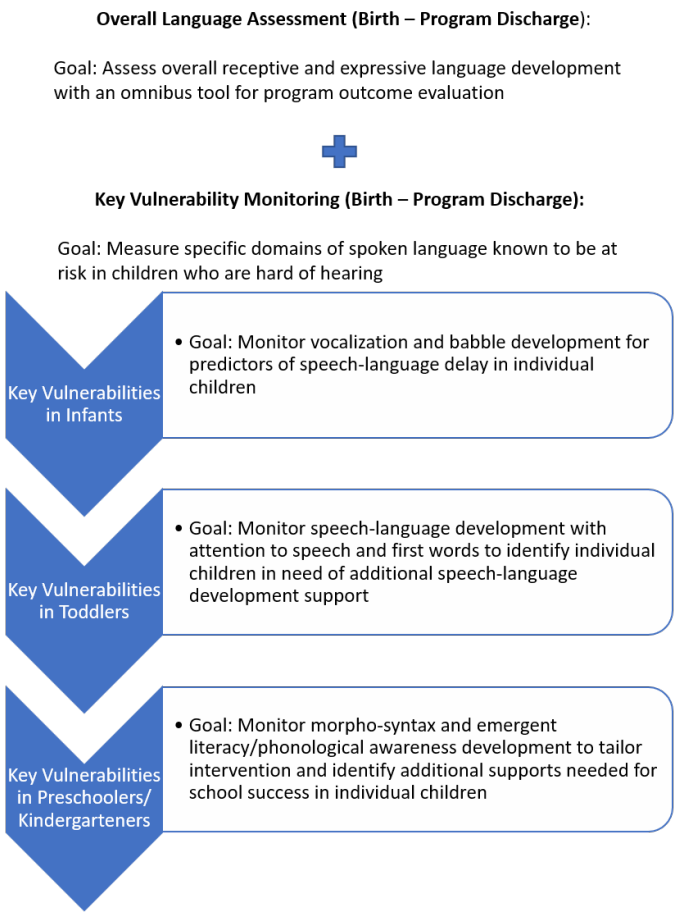
Matching the Assessment Purpose with the Assessment Method

Achieving individual level purposes requires different assessment approaches and tests than achieving program level purposes. Individual level evaluation requires different tests measuring different vulnerabilities at different stages of development. Program level evaluation requires the same metric and the same or similar tests across programs and over time. To fulfill both of these sets of purposes, it became immediately apparent that there was no single test that would be sufficient.

As a result, we suggested a two-tiered outcome monitoring framework for the IHP: (a) monitoring overall receptive and expressive language development for program-level evaluation purposes using a single test, and (b) targeted individual monitoring of selected areas of speech/language vulnerability (see Figure 2). Although we recognize that concerns in any of these domains do not clearly begin or end at any age, we recommended limiting monitoring to selected areas of speech/language vulnerability using only one or two tests at any one of three developmental time points to minimize the clinical burden of the process. This process was not intended to replace SLPs' current practices of collecting the information they need to set goals and monitor progress for individual children on their caseload. Our next step was to identify which norm-referenced tests were best equipped to measure overall expressive and receptive spoken language and each of these domains.

Figure 2

Proposed Outcome Monitoring Process



Step 2: Selecting Tests for Outcome Monitoring

Step 2a) Scoping Review of Norm-Referenced Tests

The purpose of the scoping review was to identify which norm-referenced tests have been previously used in studies of children who are HH and the results obtained using each of these tests. In developing our recommendations, we sought to select amongst tests that have a documented history of use in the peer-reviewed

literature as preliminary evidence that the tests (a) have some ability to differentiate between children who are HH and children with typical hearing thresholds and (b) are sensitive to change over time. Although the original purpose of these studies was not to document test sensitivity to group differences per se, there is a dearth of norm-referenced tests designed specifically to capture the spoken language outcomes of children who are HH. Thus, our scoping review served as our closest approximation of whether a test was likely to be sensitive enough to allow the IHP to detect group differences and change over time, should those differences or changes occur. Our expectation was that narrowing our consideration of norm-referenced assessments to only those that have been documented in the peer-reviewed literature would provide the IHP with benchmarks for spoken language outcomes, and some context to interpret their program's results. We were cognizant that if we selected a set of tests that were not sensitive to group differences, or have not previously been used with children who are HH, then we ran the risk of overestimating the outcomes of children who are HH in the IHP. Inversely, if we selected tests that were very sensitive to the spoken language vulnerabilities of children who are HH, without appropriate research context to demonstrate that these results are reasonable, we ran the risk of underestimating the outcomes of children who are HH. Although age-appropriate outcomes are appropriate goals for individual children who are HH, as a group they have been demonstrated to *statistically* perform below their same-aged peers but *within* age-expectations (e.g., Ching et al., 2013). This is not to say that EHDI programs should not strive for spoken language outcomes on par with children who are typically hearing, per JCIH recommendations (2013). However, we did not want to over- or under-estimate the IHP's impact based on artefacts of test selection.

Although EHDI intervention programs provide services to children and families electing to pursue spoken and signed language, and children who are (or are not) amplified with hearing aids or cochlear implants, our scoping review focused on articles reporting results of children who are HH who have been fitted with hearing aids and are learning a spoken language. In Ontario, cochlear implant candidacy represents a unique population who often receive services from a different publicly funded program and their outcomes are not routinely tracked by the IHP. We also restricted our review to outcomes measured in children who are HH from birth to 6 years of age to capture the language development of children who are HH in the program. Our initial review took place in 2016 across three databases (SCOPUS, CINAHL, and PubMed), but we conducted a more recent review across a modified set of databases for the purposes of this article to capture the most up-to-date publications. The results of this review were consistent with our prior review (Oram Cardy & Daub, 2017). Our review was guided by the following research questions:

1. Which tests have been used to measure spoken language in children who are HH and who have been fitted with hearing aids between birth and 6 years?
2. Which tests have been used to compare children who are HH and children with typical hearing, or subgroups of children who are HH? Which tests have detected group differences?
3. Which tests have been used to measure change over time in children who are HH? Which tests have detected change over time?

Search Strategy

Five databases were searched in October 2018: CINAHL, Pubmed, EMBASE, ERIC, and PsycInfo. Search terms were developed with the assistance of a subject librarian (see Appendix A for an example search). The search was restricted to include only studies published between 1990 and 2018 to capture research completed during the time in which the evidence supporting universal newborn hearing screening and EHDI programs began to accumulate. Following the search, the titles, abstracts, and full texts of articles were screened for several criteria. First, the article must have been published in English. Second, the article needed to have measured spoken language using a commercially available, English, norm-referenced test. Third, the study was required to report outcome data for children who are HH who wore hearing aids separately from data for children who wore cochlear implants and needed to report data for, at a minimum, a subgroup of children between birth and 6 years, 11 months. Case studies of individual children where group data was not reported were also excluded.

Title, abstract, and full text screening from articles identified through the initial database search were completed by the first author and a trained research assistant to identify articles for full review. All eligibility disagreements were resolved through discussion. Title, abstract, and full text screening from articles identified through forward and backward searching was completed by the first author using the same set of criteria previously described. This process was repeated until no new publications were identified.

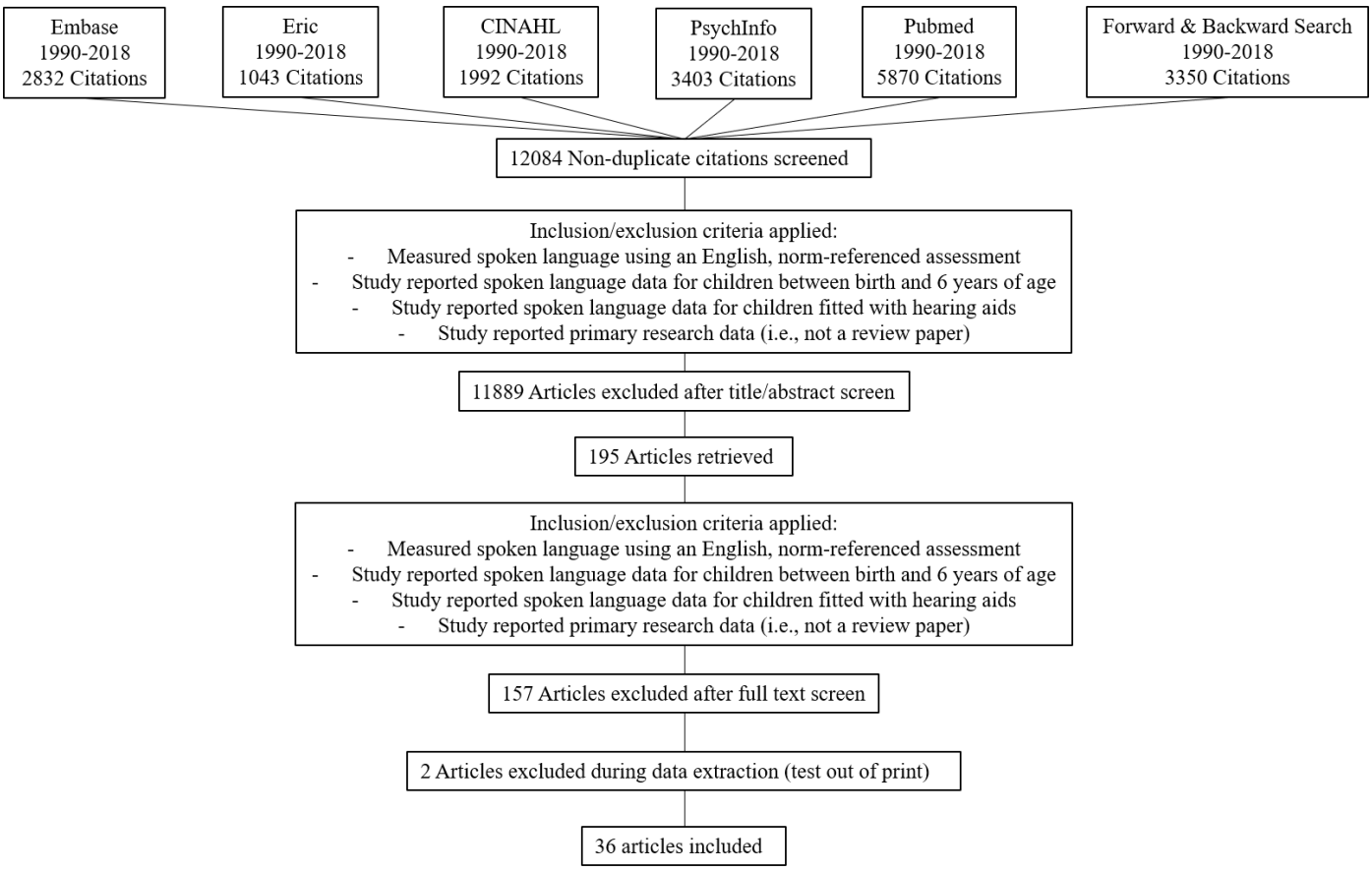
The first author extracted from each eligible article: (a) the demographic characteristics of the study population; (b) the norm-referenced test(s), including test version, used; (c) whether group comparisons were made and the results of these comparisons; and (d) whether change over time was evaluated and the results of these evaluations. At this stage, studies were excluded if the norm-referenced test was out of print (i.e., studies using only the Reynell Language Developmental Scales; Reynell & Gruber, 1990). Older versions of tests were included if there is a more recent version available for purchase. Study quality was not evaluated as the purpose of our scoping review was to capture the breadth of tools used with children who are HH and the results found with them.

Scoping Review Results

We identified 12,084 non-duplicate articles. Of those, 195 articles were retrieved after title and abstract screening. Finally, data were extracted from 36 articles (see Figure 3, and Supplemental Materials in Appendix B for the data extraction). From these 36 articles, 16 commercially available, norm-referenced tests across multiple versions were identified as having been previously used to measure spoken language outcomes in English-speaking children who are HH. Six of these tests were omnibus language measures, four were language or communication development subscales of broader developmental tests, three were measures of vocabulary, and three were measures of articulation and phonology.

Figure 3

Articles Included for Evaluation



For each test, the following was charted: the number of studies (out of 36) that used the test, whether any study used the test to make group comparisons (regardless of the results of the comparison), whether group differences were detected (out of the number of studies that used the test to evaluate group differences), whether any study used the test to measure change over time, and whether the test detected changes over time (out of the number of studies that used the test to evaluate group differences; see Table 1). Studies varied widely with respect to the ages of children included in the sample, the frequency with which they were assessed, the severity of hearing loss, characteristics of hearing aid amplification, and the demographics of comparison groups (see Supplemental Materials for further details). We identified a distinct lack of overlap in our studies in that no two studies evaluated the same outcomes in similar groups of children who are HH.

Of the 36 studies identified, 30 used 16 different norm-referenced tests to compare spoken language outcomes to other children (i.e., children with typical hearing, with cochlear implants, or with different amplification technologies) or the test's normative mean. Ten studies evaluated change over time using a variety of analyses (e.g., growth scale values, rates of language development, or linear regression). Six studies evaluated spoken language outcomes using composite scores from multiple tests using factor analyses or multivariate analyses. Only 8 out of the 16 tests were used for both comparing spoken language outcomes to other groups of children *and* measuring change over time and none of the 8 tests *consistently* identified both differences between groups and change over time.

Scoping Review Implications

The scoping review provided 16 candidate tests for measuring each of the spoken language domains within the outcome monitoring process (see Figure 2). However, one of the tests (i.e., the *Wechsler Preschool and Primary Scales of Intelligence*; Wechsler, 2002) does not primarily

Table 1

Norm-Referenced Test Use in Research with Children who are Hard of Hearing

	# of studies that used the test for any purpose	# of studies that used tests to compare groups	# of studies that found group differences	# of studies that measured change over time	# of studies that detected change over time	Of studies using composite scores (*n* = 6), # of studies using test in composite score
Omnibus language tests						
PLS	15/36	8/15	5/8	3/15	3/3	4/6
MBCDI	9/36	7/9	4[a]/7	2/9	1/2	0/6
CASL	4/36	2/4	1/2	0/4	n/a	2/6
PLAI	4/36	2/4	1/2	0/4	n/a	2/6
CELF	3/36	1/3	0/1	1/3	1/1	1/6
TACL	1/36	1/1	1[b]/1	1/1	1/1	0/6
Language scales from developmental tests						
(M)CDI	13/36	9/13	8[b]/9	0/13	n/a	2/6
VABS	5/36	2/5	0/2	0/5	n/a	2/6
MSEL	2/36	1/2	1/1	0/2	n/a	1/6
WPPSI	1/36	0/1	n/a	0/1	n/a	1/6
Vocabulary tests						
PPVT	17/36	9/17	4[b]/9	1/17	1/1	5/6
EVT	1/36	0/1	n/a	0/1	n/a	1/6
EOWPVT	2/36	1/2	1/1	1/2	1/1	0/6
Articulation/phonology tests						
GFTA	8/36	6/8	3/6	1/8	1/1	0/6
DEAP	6/36	2/6	2/2	0/6	n/a	4/6
KLPA	1/36	1/1	0/1	1/1	1[b]/1	0/6

Note. Six of the 36 reviewed studies used composite scores as an outcome measure. Multiple editions/versions of tests are combined. CASL = Comprehensive Assessment of Spoken Language; CELF = Comprehensive Evaluation of Language Fundamentals; DEAP = Diagnostic Evaluation of Articulation and Phonology; EOWPVT = Expressive One Word Vocabulary Test; EVT = Expressive Vocabulary Test; GFTA = Goldman-Fristoe Test of Articulation; KLPA = Khan-Lewis Phonological Analysis; MBCDI = MacArthur Bates Communicative Development Inventories; (M)CDI = (Minnesota) Child Development Inventory; MSEL = Mullen Scales of Early Learning; PLAI = Preschool Language Assessment Inventory; PLS = Preschool Language Scale; PPVT= Peabody Picture Vocabulary Test; TACL = Test of Auditory Comprehension of Language, VABS = Vineland Adaptive Behavior Scales.

measure spoken language, and largely measures domains that fall outside SLPs' scope of practice in the province of Ontario. Therefore, it was excluded from future evaluations. Additionally, the *Expressive Vocabulary Test* (Williams, 2007) was used once in previous studies as a part of a composite score and was not used in studies making group comparisons or evaluating change over time. Given the lack of data about the *Expressive Vocabulary Test's* performance on its own, we excluded it from future evaluations. Our next step was to examine the psychometric properties of each of the 14 candidate tests to determine which ones would be psychometrically appropriate to meet the IHP's assessment purposes.

Step 2b) Critical Appraisal of Norm-Referenced Tests

After completing the initial 2016 scoping review, the most recent versions of the 14 tests, regardless of whether they were the versions used in studies included in the scoping review, were evaluated using the 2012 version of the Consensus Based Standards for the Selection of Health Status Measurement Instruments (COSMIN; Mokkink et al., 2012) checklist. The COSMIN checklist was developed using an International Delphi study method where experts in fields related to measurement (e.g., epidemiology and statistics) iteratively responded to a series of questions about which measurement properties ought to be evaluated in test design (specifically Health-Related Patient Reported Outcomes, but with application to other tests) and the statistics that should be used to report them. Consensus (greater than 67% agreement) was reached on most major terms (with the exception of structural validity), definitions of each property, and on the taxonomy's organization. From this taxonomy, the COSMIN team developed quality criteria for both

the methodological quality of studies designed to collect data information about measurement properties, and the measurement properties themselves (Terwee, 2011). For the purposes of developing our recommendations, we focused our evaluation on the quality of the measurement properties reported in the examiner's manual, but not the methodological quality of the studies designed to report the measurement properties, as it was quite likely that not all examiner's manuals would report sufficient detail to adequately appraise the quality of the methods themselves.

Critical Appraisal Analysis

To appraise each test, we used a revised version of the COSMIN quality criteria in which we excluded four criteria that were included in the original checklist (criterion validity, cross-cultural validity, responsiveness, and measurement error). Although we agree that these criteria are important to consider, upon review it became clear that the statistics required to evaluate these criteria (e.g., differential item functioning analyses between multiple language versions) were very rarely evaluated in any of the included tests, and evaluating these criteria would not support us in choosing a test amongst the 14 tests we identified. Therefore, each of the 14 tests were appraised with respect to the following: internal consistency, reliability, content validity, construct validity (hypothesis testing), and construct validity (structure). Each domain was assigned one of three ratings (positive, indeterminate, negative) according to the operationalizations of each criterion in the COSMIN checklist. For example, a test was rated as having positive evidence for structural validity if factors explained 50% or more of the variance, indeterminate if explained variance was not evaluated/discussed, or negative if factors explained 49% or less of the variance. For our purposes, we considered a test to have met reasonable criteria if they received a positive rating in at least 4 of the 5 categories.

Critical Appraisal Results

Only eight of the 14 tests met acceptable criteria in 4 of the 5 appraised COSMIN domains (see Table 2). Within each of the test categories (omnibus/language scale, vocabulary, phonology/articulation; Table 2), at least one test met acceptable criteria in 4 of the 5 COSMIN domains. Most tests (12 of the 14) met acceptable criteria for reliability, and all tests reported at least one measure of reliability. Only one test reported weak evidence for validity domains, but most tests were missing validity information. Information about tests' internal structure was the least frequently reported (only two of the 14 tests) in examiner's manuals.

Critical Appraisal Implications

Based on our appraisal, we identified eight norm-referenced tests that were largely psychometrically acceptable to select for the spoken language outcome monitoring process. There was not one test with clearly better measurement properties over the others. Our next step was to summarize the administration properties of each of these tests.

Step 2c) Consideration of Administration Properties

We considered various administration properties in summarizing the candidate tests including: the age ranges for which each test had normative data; whether the test covered overall language abilities or subskills; the types of scores that could be calculated (e.g., percentile ranks and/or growth scale values), who was required to administer the test (clinician or caregiver), and the amount of time each test took to administer. Each of the eight acceptable tools had various administration properties that might make the test more, or less, attractive to individual EHDI programs (Table 3). For instance, the PLS-5, *Clinical Evaluation of Language Fundamentals, Preschool, 2nd ed.* (CELF-P2; Semel et al., 2004) and *Comprehensive Assessment of Spoken Language, 2nd ed.* (CASL-2; Carrow-Woolfolk, 2017) were all acceptable omnibus language measures, but the PLS-5 provides scores that support measuring change over time (i.e., growth scale values), the CELF-P2 supports profiling different domains of language, and the CASL-2 measures a broader range of language abilities and is appropriate at older ages than either the PLS-5 or CELF-P2. Therefore, consideration of these properties presented us with flexibility in which test(s) to propose. For the purpose of the IHP, tests like the PLS-5 had administration properties that would enable the IHP to achieve more of their outcome monitoring purposes. Specifically, the PLS-5 reported normative data for all age ranges served by the program and also reported growth scale values, which would enrich program level evaluation of growth over time. However, other tests had other relative advantages over the PLS-5. For instance, the *MacArthur-Bates Communicative Development Inventories, 2nd ed.* (MBCDI-2; Fenson et al., 2007) could be completed by parents without SLPs' support, and the CELF-P2 supported profiling. Our next step was to triangulate the administrative properties and relative advantage of each test with the evidence for the quality of each test to develop a set of options. We then shared these initial recommendations with the IHP and a panel of expert SLPs who had volunteered their time to provide feedback on the clinical feasibility of our recommendations.

Step 3) Integrating the Evidence into Recommendations

Recommendations for Overall Spoken Language Outcome Monitoring

In accordance with JCIH recommendations, we proposed that all children in the IHP be tested with a standardized measure that compares their spoken language development to that of same-aged children with typical hearing every 6 months during the first 3 years of life, and every year thereafter. Triangulation of the evidence from our scoping review, critical appraisal, and summary of administration properties indicated that the following three measures had the strongest evidence supporting their selection as a measure of overall language abilities: PLS-5, MBCDI-2, and CELF-P2. Both the PLS-5 and CELF-P2 offer the additional advantages of having diagnostic accuracy information with cut-point scores

Table 2
Critical Appraisal of Norm-Referenced Tests Using COSMIN Criteria

	Internal Consistency	Reliability	Content Validity	Hypothesis Testing	Structure
Omnibus language tests					
PLS-5	+	+	+	+	?
MBCDI-2	+/-	+	+	+	?
CASL-2	+	+	+	+	?
PLAI-2	?	+/-	?	+/-	+
CELF-P2	+/-	+	+	+	+
TACL-4	+	?	?	+	?
Language scales from developmental tests					
CDI	+	+	+	?	?
MSEL	?	+/-	?	+	?
VABS-3	+	+/-	?	+	?
Vocabulary tests					
PPVT-4	+	+	+	+	?
EOWPVT-4	+	+	+	-	?
Articulation/phonology tests					
GFTA-3	+	+	+	+	?
DEAP	+	+/-	+	+/-	?
KLPA-3	+	+	+	+	?

Note. Ratings included positive evidence (+), indeterminate (?), and negative evidence (-) in meeting COSMIN Criteria. +/- indicates that some, but not all, subtests meet acceptable criteria. Shaded tests received a positive rating in at least 4/5 of the categories. Preschool Language Scale (PLS-5; Zimmerman et al., 2011); MacArthur Bates Communicative Development Inventories (MBCDI-2; Fenson et al., 2007); Comprehensive Assessment of Spoken Language (CASL-2; Carrow-Woolfolk, 2017); Preschool Language Assessment Inventory (PLAI-2; Blank et al., 2003); Comprehensive Evaluation of Language Fundamentals (CELF-P2; Semel et al., 2004); Test of Auditory Comprehension of Language (TACL-4; Carrow-Woolfolk, 2014); Child Development Inventory (CDI; Ireton, 1992); Mullen Scales of Early Learning (MSEL; Mullen, 1995); Vineland Adaptive Behavior Scales (VABS-3; Sparrow et al., 2016); Peabody Picture Vocabulary Test (PPVT-4; Dunn & Dunn, 2007); Expressive One Word Vocabulary Test (EOWPVT-4; Martin & Bronwell, 2011); Goldman-Fristoe Test of Articulation (GFTA-3; Goldman & Fristoe, 2015); Diagnostic Evaluation of Articulation and Phonology (DEAP; Dodd et al., 2006); Khan-Lewis Phonological Analysis (KLPA-3; Khan & Lewis, 2002).

and growth scale values. The PLS-5 covers the full 0 to 6 year age range serviced by the IHP, while the CELF-P2 covers 3 to 6 years, and the MBCDI includes three separate forms that cover 8 to 18 months (MBCDI Words and Gestures), 16 to 30 months (MBCDI Words and Sentences), and 30 to 37 months (MBCDI III). Therefore, the most parsimonious approach would be to use the PLS-5 across the entire age span of the program. However, we have encountered SLPs and scientific experts in the field of permanent childhood hearing loss (e.g., Dr. Mary Pat Moeller, personal communication) who have expressed concerns about the sensitivity of the PLS-5 in the first two years of life. These concerns are consistent with the diagnostic accuracy data reported in the examiner's manual (Zimmerman et al., 2011). That is, the PLS-5's diagnostic accuracy does not meet acceptable criterion (≥ 0.80; Plante & Vance, 1994) for detecting language delays in children under 2 years for any cut-score. Therefore, although using the PLS-5 would allow the IHP to evaluate whether children were making significant progress over time, SLPs would be unable to accurately determine whether children were obtaining age-appropriate outcomes *and* the PLS-5 posed greater clinical burden (i.e., longer administration time) than other candidate tests.

An alternative option could be to use the three separate forms of the MBCDI-2 in the first three years of life and the CELF-P2 thereafter. However, because the subtests and scores on the three MBCDI-2 forms are different, this would prohibit future analysis of developmental growth over time, which "can only be analyzed if the child is assessed with at least some instruments that can be

Table 3

Administration Properties for Currently Available Versions of Psychometrically Suitable Norm-Referenced Tests

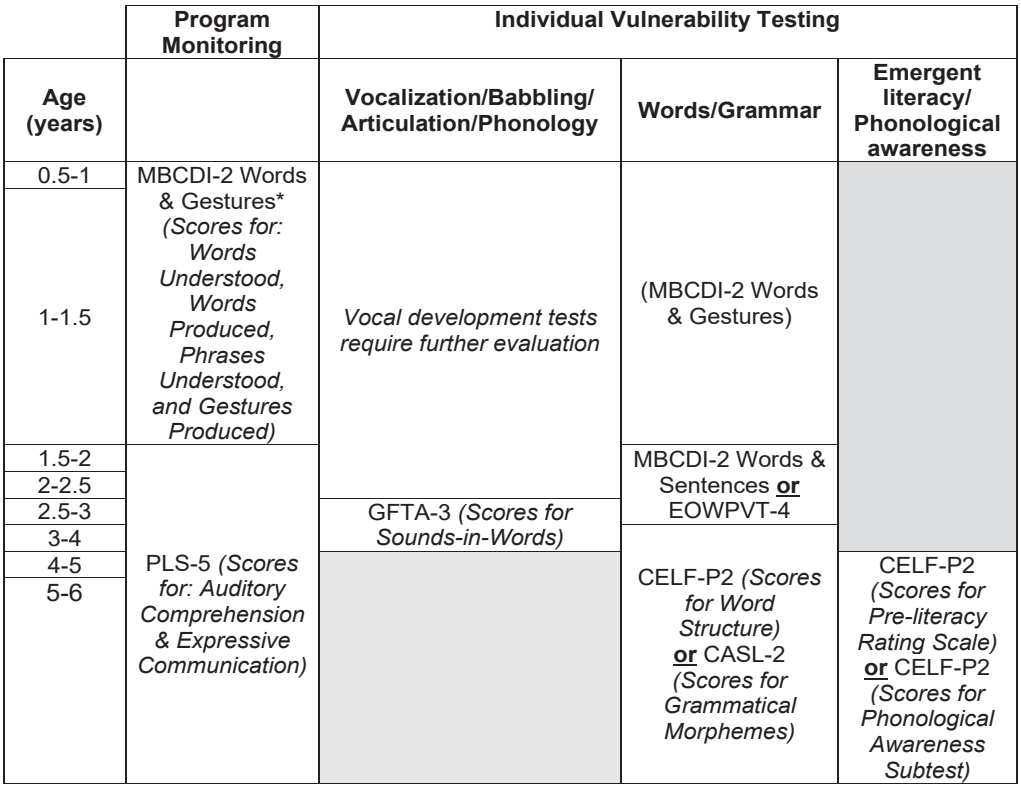
	Age range	Language Areas		Scores Available				Examiner/ Respondent	Time (min)
		Overall	Subskills	SS	GSV	PR	AE		
PLS-5	0-7 years	✓		✓	✓	✓	✓	Clinician	45–60
MBCDI-2	8-18, 16-30, 30-37 months		✓			✓		Caregiver	20–40
CASL-2	3-6 years	✓	✓	✓		✓	✓	Clinician	30–45
CELF-P2	3-6 years	✓	✓	✓	✓	✓	✓	Clinician	varies
PPVT-4	2;6-90 years		✓	✓	✓	✓	✓	Clinician	8–16
GFTA-3	2-21 years		✓	✓	✓	✓	✓	Clinician	5–10
DEAP	3-8 years		✓	✓			✓	Clinician	5–15
KLPA-3	8-21 years		✓	✓		✓	✓	Clinician	10–30

Note. AE = age equivalent; GSV = gross scale value; PR = percentile rank; SS = standard score.

Preschool Language Scale (PLS-5; Zimmerman et al., 2011); MacArthur Bates Communicative Development Inventories (MBCDI-2; Fenson et al., 2007); Comprehensive Assessment of Spoken Language (CASL-2; Carrow-Woolfolk, 2017); Comprehensive Evaluation of Language Fundamentals (CELF-P2; Semel et al., 2004); Peabody Picture Vocabulary Test (PPVT-4; Dunn & Dunn, 2007); Goldman-Fristoe Test of Articulation (GFTA-3; Goldman & Fristoe, 2015); Diagnostic Evaluation of Articulation and Phonology (DEAP; Dodd et al., 2006); Khan-Lewis Phonological Analysis (KLPA-3; Khan & Lewis, 2002).

Figure 4

Final Recommendation



Note. CASL-2 = Comprehensive Assessment of Spoken Language; CELF-P2 = Comprehensive Evaluation of Language Fundamentals; EOWPVT = Expressive One Word Vocabulary Test; GFTA-3 = Goldman-Fristoe Test of Articulation; MBCDI-2 = MacArthur Bates Communicative Development Inventories; PLS-5 = Preschool Language Scale.

repeated throughout the target age range" (JCIH, 2013, p. e1334). An additional concern is that only the MBCDI Words and Gestures form includes evaluation of both receptive and expressive language (along with gestures); the remaining MBCDI-2 forms only assess expressive language.

A third option included using the MBCDI-2 Words and Gestures form until 18 months of age, and the PLS-5 thereafter. This would provide scores on the same measure (the MBCDI-2) for the first two testing sessions at the 6-month testing interval, and then PLS-5 scores for all 6-month and 12-month testing intervals beyond 18 months. Under this option, the program would be able to make direct comparisons of growth across all time points except for the one point of transition between the MBCDI-2 and PLS-5 around 18 to 24 months. We felt that this was a reasonable compromise to have a more clinically accepted tool in the earliest years of development, and thus this third option formed the basis for our final recommendation.

Recommendations for Individual Vulnerability Testing

Our scoping review and critical appraisal identified norm-referenced tests that have been used with children who are HH and that measure areas that are particularly vulnerable for them. Based on the results of our scoping review and critical appraisal, we recommended a two-pronged approach to assessment for the purposes of supporting individual child/family needs. We recommended that SLPs include assessment of key vulnerabilities associated with the child's particular age/stage of development (see Figure 4) alongside of their administration of the program-level test of overall language abilities. To reduce the time associated with assessment, and to prevent children from being assessed with more than two norm-referenced tests at a single session, we recommended assessing one area of key vulnerability at each age, even though the ages at which different skills (e.g., articulation and phonology) can be assessed may overlap with other key vulnerabilities. Additionally, in our scoping review we were unable to identify any commercially available test of early vocal development, although some articles (e.g., Ambrose et al., 2014) report on experimental tests that are currently in development. In this regard, we were unable to recommend a specific test for the IHP to use for monitoring early vocal development. In short, we recommended that the IHP provide a set of recommended tests from which SLPs are advised to select. This would support consistency across regions and ensure that only those tests with the strongest evidence are used to assess these key vulnerability areas.

Consultation with Stakeholders

We summarized the overall process (program level monitoring and individual vulnerability testing) as well as the three options for overall outcome monitoring and our recommendations for individual vulnerability testing (described above), in a formal written report (Oram Cardy & Daub, 2017). This report was shared with IHP audiological policy development, IHP government leaders, and a team of SLPs who formed an advisory panel. All parties provided written feedback on the report and discussed the recommendations at length through teleconference meetings. Following the revisions to the recommendations, all parties reached agreement on a final procedure (see Figure 4). This procedure included program-level outcome monitoring and individual vulnerability testing. Following final discussion via teleconference, the managerial team ultimately adopted the final spoken language outcome monitoring procedure for implementation in the IHP.

Discussion

The present article describes our process for developing a set of spoken language outcome monitoring recommendations to support a Canadian EHDI program, the Ontario IHP, in fulfilling best practice recommendations. To date, there has been limited guidance in the literature on (a) the best way to approach the development of a spoken language outcome monitoring process or (b) how to accomplish all of the facets of spoken language outcome monitoring in a way that provides statistically appropriate evidence, is implementable across entire EHDI programs, and meets the competing needs of different stakeholders. Our expectation is that documenting our steps in this process and the recommendations that resulted will not only provide a general framework and example for other EHDI programs, but also highlight the previously undiscussed challenges of designing such a procedure.

Our process was grounded in the initial JCIH (2007, 2013) recommendations for spoken language as well as consideration of the International Consensus work on best practice principles (Moeller et al., 2013). From this foundation, we considered the purposes of spoken language outcome monitoring from the perspective of various IHP stakeholders to clarify the assessment purposes our process would need to fulfill. Using these purposes, we conducted a scoping review to identify a set of candidate norm-referenced tests that have been previously used to fulfill these assessment purposes and appraised the psychometric quality of the most recent versions of these tests. We then considered the administration properties of the tests that we rated as psychometrically acceptable and integrated all sources of evidence with our originally described assessment purposes. This allowed us to develop a set of recommendations to share with IHP stakeholders, who ultimately decided to adopt them. We expect that our work will be of interest to other EHDI programs and service providers who work with children who are DHH by documenting our process in developing our recommendations, the recommendations themselves, and the final procedure adopted by the IHP. Our results highlight the unique challenges faced when trying to develop a process for spoken language outcome monitoring, guide future research designed to refine the development process, and contribute to a body of literature that provides guidance for EHDI programs looking to fulfill best practice recommendations.

Our next step is to design implementation materials and conduct pilot projects to evaluate the new procedures for both overall spoken language monitoring and individual vulnerability monitoring. These pilot projects are intended to identify barriers and facilitators to implementing the new recommendations in clinical practice, and to allow us to refine our process into one that is most sustainable and clinically feasible before program-wide launch. We anticipate that the results of these pilot projects will similarly support discussions of spoken language outcome monitoring in EHDI programs and highlight the inherent complexity in accomplishing these goals.

We do not intend to assert that our process or final recommendations are a gold standard for spoken language outcome monitoring and should be adopted by other EHDI programs. Rather, we believe that our work uniquely highlights the challenges in accomplishing spoken language outcome monitoring and may be a valuable foundation for EHDI programs looking to develop, or refine, their spoken language outcome monitoring procedures. Our projects were developed through the lens of the Ontario IHP, and other EHDI programs might have different priorities for spoken language outcome monitoring, amongst other needs. In our case, the IHP sought a process that would allow them to use the data to evaluate whether children across the province are making progress in their spoken language over time, whether they are meeting age-appropriate expectations by the time they are discharged from the program, and whether they have the spoken language skills they need at discharge to be prepared for school. Necessarily, fulfilling these purposes required the use of multiple tests that are sensitive to multiple domains of language, and that were norm-referenced to establish whether a child was performing within or below age-expectations.

An additional priority was selecting norm-referenced tests from those that have been previously used in research with children who are HH to contextualize the outcomes in the IHP with the peer-reviewed literature. The Ontario IHP is publicly funded and managed under a larger provincial division also responsible for the allocation of resources across multiple programs from a single budget. We were wary of selecting norm-referenced tests without a documented history of use in the literature because it has been demonstrated that children who are HH often score within age-expectations (and close to the test's normative mean of a standard score of 100), but statistically lower than matched groups of children with typical hearing (e.g., Tomblin et al., 2015). In this case, using a standard score cut-off recommended by a norm-referenced test was not sufficient to describe program outcomes. We were aware that spoken language outcome data could be used by policy makers to make funding decisions and that there was a risk of misinterpreting program level outcomes as being insufficient to continue funding. We were also aware that EHDI programs are precariously positioned in Canada: many EHDI programs are in development, and some have seen declines in support from previous years (Canadian Infant Hearing Task Force, 2014; 2019). In the Canadian context, statistically sound outcome data from one EHDI program has the potential to provide evidence to influence other provincial or national funding priorities. Therefore, it was critical to develop a process that we could connect to the peer-reviewed literature to evaluate whether the IHP was performing on par with documented outcomes in other EHDI programs.

Even within the context of the Ontario IHP, our recommendations remain limited in a number of respects. Canada has two official languages (English and French) and many regions in the province are densely populated, multicultural areas where residents speak languages other than these. We focused our reviews and recommendations on measuring outcomes for children who are HH from English speaking families, in part, due to a dearth of norm-referenced tests that have been validated in other languages to include in our scoping review and critical appraisal. Certainly, many (but not all, i.e., the MBCDI-2) of the tests we selected for our current recommendations have not been normed in French, even if there are translated versions (i.e., the PLS-5). To fulfil clinical assessment needs, we have advised SLPs to continue using the tools they typically would for children for whom English is not a primary language, although their outcomes will not be able to be evaluated at the program-level in the provincial database. This raises concerns about equitable service provision—regardless of the language their child is learning, families deserve to know whether their child is progressing as expected in response to intervention. Solutions and next steps, such as collecting local normative data on translated versions, are under discussion. Until norm-referenced assessments for these groups of children exist, EHDI programs will need to identify other creative solutions to evaluate spoken language outcomes and rely on less formal assessments. Our general framework could be modified to support identifying informal assessments or interview tools, although a different process for critically appraising the approaches would be needed.

It is likely that there are other important considerations requiring attention in other EHDI programs that we did not account for in our process for the Ontario IHP. For example, EHDI programs in which outcome data are not likely to be used to support funding decisions may feel comfortable considering the use of norm-referenced tests without a history of previous peer-reviewed use. Additionally, our process did not consider the spoken language outcomes of children with cochlear implants because many are served by a different program in the province of Ontario, but other EHDI programs may wish to do so. Furthermore, our process did not attend to the sensitivity and specificity cut-off scores for language impairment on the tests we evaluated because there is no mandate in Ontario for children to perform below a certain threshold (e.g., -2 *SD* below the mean) to be considered eligible for receiving SLP services outside of EHDI programming. This is certainly the case in some American state education departments (Spaulding et al., 2012), thus, EHDI programs located in regions with similar

requirements will need to additionally consider whether candidate tests are adequately sensitive/specific at the cut-off scores required to receive services.

Despite these limitations, our experience has highlighted major challenges in fulfilling spoken language outcome monitoring worthy of further consideration by the field. There is certainly more room for discussion about which assessment considerations ought to be prioritized in developing spoken language outcome monitoring procedures, the role of norm-referenced tests versus other sources of assessment information (e.g., criterion referenced testing for goal setting), and ways to ensure equity in how these sources of information are collected and used across programs. First, outcomes from two norm-referenced tests are not directly comparable and the operationalization of "within age-expectations" is entirely dependent on the statistical properties of the norm-referenced test in question. Although the JCIH recommends that children who are HH should score within -1 *SD* of the mean or higher on norm-referenced tests (2013), this recommendation does not acknowledge the unique sensitivity and specificity of individual tests at individual scores (Spaulding et al., 2006). For example, both the PLS-5 and the CELF-P2 have the greatest diagnostic accuracy at -1 *SD* (Zimmerman et al., 2011; Semel et al., 2004), but the GFTA-3 maximizes diagnostic accuracy at -1.5 *SD* (Goldman & Fristoe, 2015). As such, children with typical hearing thresholds and typical language development can be expected to score between -1.49 and -1 *SD*s below the mean on the GFTA-3. If stakeholders apply the -1 *SD* cut-off as the expectation on tests that are less accurate at -1 *SD*, they may be inadvertently holding children who are HH to a higher standard than their peers with typically developing hearing. In other words, defining age-appropriate outcomes for individual children, and appropriate outcomes for children who are HH as a group, is confounded with the psychometric properties of norm-referenced tests (Spaulding et al., 2006). These confounds pose significant challenges to stakeholders looking to interpret their population level outcome data. A program that elects to use the PLS-5 to measure outcomes might *appear* to have better outcomes (i.e., within -1 *SD* of the mean) than a program that elects to use a test with a -1.5 *SD* cut-off, even though the children in both programs might be performing within age-expectations. Therefore, procedures for measuring outcomes *must* consider the unique psychometric properties of the tests they are using or risk generating data that suggests their program is failing to meet JCIH benchmarks.

These concerns with defining age-appropriate outcomes and interpreting results are compounded when we consider applying spoken language outcome monitoring to different groups of children, including those 20% to 40% of children who are HH who have additional diagnoses, some of which (e.g., autism, cerebral palsy, and developmental delay) may further impact language development (Cupples, Ching, Crowe, Day, et al., 2014). Future work could extend the methods used here to identify studies examining language outcomes in children with an additional diagnosis, with and without hearing loss. This would provide context to any program looking to report on the results of children who are HH with additional disabilities.

A second challenge with accomplishing spoken language outcome monitoring pertains to the clinical feasibility of accomplishing all necessary assessment purposes. Many norm-referenced tests are not developed to serve multiple assessment purposes, and their use is best restricted to interpreting whether a child is, or is not, within age-expectations. This creates challenges for accomplishing the diverse purposes that spoken language outcome monitoring is intended to fulfill (e.g., treatment planning and evaluating EHDI programs broadly). Some of these purposes can certainly be accomplished through other forms of assessment (e.g., criterion referenced assessment, language sample analysis), and neither we, nor the JCIH (2013), argue that norm-referenced assessments should be the *only* component of a spoken-language outcome monitoring battery. Certainly, SLPs will need to rely on other sources of information to develop their therapy plans. However, the addition of a standard norm-referenced process to fulfill program-level evaluation goals adds lengthy tasks to SLPs' assessment time and it is unknown whether it is feasible for SLPs to collect, interpret and integrate all of the necessary sources of information needed to fulfill spoken language outcome monitoring recommendations. It is widely accepted that whether research evidence or new recommendations will be successfully used in clinical practice is influenced by numerous factors within the clinical context (e.g., Dobrow et al., 2004; Graham et al., 2006) such as time, caseload, and clinician factors (e.g., beliefs, knowledge, skills) above and beyond the quality of the research evidence or recommendation itself. Accomplishing spoken language outcome monitoring in EHDI programs is complicated not only by limited evidence to guide development of procedures, but also by a lack of evidence to support implementation of these procedures. To our knowledge, there is only one peer-reviewed paper, published by our research group (Cunningham et al., 2019) that has evaluated SLPs' perceptions of the barriers to implementing spoken language outcome monitoring in an EHDI program. In Cunningham's investigation, time for additional testing was a primary concern. Additional work is needed to evaluate the feasibility of our recommendations specifically, and spoken language outcome monitoring broadly, as well as to develop implementation interventions that result in effective, sustained uptake of spoken language outcome monitoring procedures.

Conclusion

Guidance for how to best implement spoken language outcome monitoring recommendations (JCIH 2007; 2013) is lacking, and EHDI programs face significant barriers to developing procedures that fulfill best-practice recommendations. The present article describes a series of projects, conducted as part of program evaluation and

quality improvement for the Ontario IHP, to develop a spoken language outcome monitoring procedure using a scoping review and critical appraisal of candidate norm-referenced tests. We expect that the process we used, the recommendations we developed, and the challenges we encountered, will be informative to other EHDI programs looking to develop their own procedures. Final recommendations included developing a two-tiered assessment battery measuring overall spoken language outcomes and key areas of spoken language vulnerability. Future work evaluating the appropriateness of these recommendations, whether the data collected is sufficient to fulfill our intended purposes, the feasibility of our recommendations and ways to implement them into clinical practice are needed.

References

Ambrose, S. E., Unflat Berry, L. M., Walker, E. A., Harrison, M., Oleson, J., & Moeller, M. P. (2014). Speech sound production in 2-year-olds who are hard of hearing. *American Journal of Speech-Language Pathology*, *23*(2), 91–104. https://doi.org/10.1044/2014_AJSLP-13-0039

Blank, M., Rose, S. A., & Berlin, L. J. (2003). *Preschool language assessment instrument – Second edition (PLAI-2)*. Pro-ed.

Canadian Infant Hearing Task Force. (2014). Report card of early hearing detection and intervention. https://www.sac-oac.ca/sites/default/files/resources/Report%20Card-2014_EN.pdf

Canadian Infant Hearing Task Force. (2019). 2019 Report card on Canadian early hearing detection and intervention programs. https://canadianaudiology.ca/wp-content/uploads/2019/04/Report-Card-2019_FINAL_VERSION_EN.pdf

Carrow-Woolfolk, E. (2014). *Test for Auditory Comprehension of Language, 4th ed.* Pro-ed.

Carrow-Woolfolk, E. (2017). *Comprehensive Assessment of Spoken Language, 2nd ed.* Pearson Assessments.

Ching, T. Y. C., Day, J., Dillon, H., Gardner-Berry, K., Hou, S., Seeto, M., Wong, A., & Zhang, V. (2013). Impact of the presence of auditory neuropathy spectrum disorder (ANSD) on outcomes of children at three years of age. *International Journal of Audiology*, *52*(sup2), S55–S64. https://doi.org/10.3109/14992027.2013.796532

Ching, T. Y. C., Dillon, H., Button, L., Seeto, M., Van Buynder, P., Marnane, V., Cupples, L., & Leigh, G. (2017). Age at intervention for permanent hearing loss and 5-year language outcomes. *Pediatrics*, *140*(3), e20164274. https://doi.org/10.1542/peds.2016-4274

Cunningham, B. J., Daub, O., & Oram Cardy, J. (2019). Barriers to implementing evidence-based assessment procedures: Perspectives from the front lines in speech-language pathology. *Journal of Communication Disorders*, *80*, 66–80. https://doi.org/10.1016/j.jcomdis.2019.05.001

Cupples, L., Ching, T. Y. C., Crowe, K., Day, J., & Seeto, M. (2014). Predictors of early reading skill in 5-year-old children with hearing loss who use spoken language. *Reading Research Quarterly*, *49*(1), 85–104. https://doi.org/10.1002/rrq.60

Daub, O. (2016). Predictors of language outcome for children in the Ontario Infant Hearing Program. [Unpublished master's thesis]. University of Western Ontario, London ON.

Daub, O., Bagatto, M. P., Johnson, A. M., & Oram Cardy, J. (2017). Language outcomes in children who are deaf and hard of hearing: The role of language ability before hearing aid intervention. *Journal of Speech, Language, and Hearing Research*, *60*(11), 3310–3320. https://doi.org/10.1044/2017_JSLHR-L-16-0222

Daub, O., Cunningham, B. J., Bagatto, M. P., Johnson, A. M., Kwok, E. Y. L., Smyth, R. E., Oram Cardy, J. (in press). Adopting a conceptual validity framework for testing in speech-language pathology. *American Journal of Speech-Language Pathology.*

Daub, O., Skarakis-Doyle, E., Bagatto, M. P., Johnson, A. M., & Oram Cardy, J. (2019). A comment on test validation: The importance of the clinical perspective. *American Journal of Speech-Language Pathology*, *28*(1), 204–210.

Dobrow, M. J., Goel, V., & Upshur, R. E. G. (2004). Evidence-based health policy: Context and utilisation. *Social Science & Medicine, 58*, 207–217.

Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2006). *Diagnostic Evaluation of Articulation and Phonology (DEAP)*. Pearson Education Inc.

Dunn, L. M., Dunn, D. M. (2007). *Peabody Picture Vocabulary Test, 4th ed.* Pearson Education Inc.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., Bates, E. (2007). *The MacArthur-Bates Communicative Development Inventories User's Guide and Technical Manual 2nd edition.* Brookes Publishing.

Goldman, R., & Fristoe, M. (2015). *Goldman-Fristoe Test of Articulation, 3rd ed.* Pearson Education Inc.

Graham, I. D., Logan, J., Harrison, M. B., Straus, S. E., Tetroe, J., Caswell, W., & Robinson, N. (2006). Lost in knowledge translation: Time for a map? *The Journal of Continuing Education in the Health Professions, 26*, 13–24.

Ireton, H. (1992). *Child Development Inventory, Manual*. Behavior Science Systems.

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. (2013). Supplement to the JCIH 2007 position statement: Principles and guidelines for early intervention after confirmation that a child is deaf or hard of hearing. *Pediatrics, 131*(4), e1324–e1349.

Joint Committee on Infant Hearing. (2019). Year 2019 position statement: Principles and guidelines for early hearing detection and intervention programs. *Journal of Early Hearing Detection and Intervention, 4*(2), 1–44. https://doi.org/10.15142/fptk-b748

Khan, L., & Lewis, N. P. (2002). *Khan-Lewis Phonological Analysis 2nd edition.* Pearson Education Inc.

Martin, N. A., & Brownell, R. (2011). *Expressive One-Word Picture Vocabulary Test – 4th edition.* Pro-ed.

McCauley, R. R., & Swisher, L. (1984). Use and misuse of norm-referenced tests in clinical assessment: A hypothetical case. *Journal of Speech and Hearing Disorders, 49*, 338–348.

Moeller, M. P. (2000). Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics*, *106*(3), e43–e51. https://doi.org/10.1542/peds.106.3.e43

Moeller, M. P., Carr, G., Seaver, L., Stredler-Brown, A., & Holzinger, D. (2013). Best practices in family-centered early intervention for children who are deaf or hard of hearing: An international consensus statement. *Journal of Deaf Studies and Deaf Education*, *18*(4), 429–445. doi.org/10.1093/deafed/ent034

Moeller, M. P., Hoover, B. M., Putman, C. A., Arbataitis, K., Bohnenkamp, G., Peterson, B., Wood, S., Lewis, D., Pittman, A., & Stelmachowicz, P. (2007). Vocalizations of infants with hearing loss compared to infants with normal hearing. Part I: Phonetic development. *Ear and Hearing, 28*(5), 605–627.

Moeller, M. P., Hoover, B. M., Putman, C. A., Arbataitis, K., Bohnenkamp, G., Peterson, B., Lewis, D., Estee, S., Pittman, A., Stelmachowicz, P. (2007). Vocalizations of infants with hearing loss compared to infants with normal hearing. Part II: Transition to words. *Ear and Hearing, 28*(5), 628–642.

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), 740–753. doi.org/10.1097/AUD.0b013e318157f07f

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2012). The COSMIN checklist manual, *VU University Medical,* 1–56. https://fac.ksu.edu.sa/sites/default/files/cosmin_checklist_manual_v9.pdf

Mullen, E. M. (1995). *Mullen Scales of Early Learning* (AGS ed.). American Guidance Service Inc.

Oram Cardy, J., & Daub, O. (2017). Recommendations for the assessment of spoken language development in the Ontario Infant Hearing Program. Report to the Ontario Ministry of Children, Community, and Social Services.

Peña, E. D., Spaulding, T. J., & Plante, E. (2006). The composition of normative groups and diagnostic decision making: Shooting ourselves in the foot. *American Journal of Speech-Language Pathology*, *15*, 247–254. https://doi.org/10.1044/1058-0360(2006/023)

Plante, E., & Vance, R. (1994). Selection of preschool language tests: A data-based approach. *Language, Speech, and Hearing Services in Schools*, *23*, 15–24. https://doi.org/10.1044/0161-1461.2501.15

Reynell, J., & Gruber, C. (1990). *Reynell Developmental Language Scales*. Western Psychological Services.

Semel, E., Wiig, E., & Secord, W. A. (2004). *Clinical Evaluation of Language Fundamentals – Preschool, 2nd ed.* Pearson Education Inc.

Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland Adaptive Behavior Scales, 3rd edition.* Pearson Education Inc.

Spaulding, T. J., Plante, E., & Farinella, K. A. (2006). Eligibility criteria for language impairment: Is the low end of normal always appropriate? *Language Speech and Hearing Services in Schools, 37*(1), 61–72.

Spaulding, T. J., Szulga, M. S., & Figueroa, C. (2012). Using norm-referenced tests to determine severity of language impairment in children: Disconnect between U.S. policy makers and test developers. *Language, Speech, and Hearing Services in Schools, 43*, 176–190. https://doi.org/10.1044/0161-1461(2011/10-0103)

Terwee, C. B. (2011). Protocol for systematic reviews of measurement properties. In *Knowledge Center Measurement Instruments*. Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, Amsterdam: The Netherlands.

Tomblin, J. B., Harrison, M., Ambrose, S. E., Walker, E. A., Oleson, J. J., & Moeller, M. P. (2015). Language outcomes in young children with mild to severe hearing loss: *Ear and Hearing*, *36*, 76S–91S. https://doi.org/10.1097/AUD.0000000000000219

Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence- Third Edition.* The Psychological Corporation.

Williams, K. T. (2007). *Expressive Vocabulary Test Second Edition*. Pearson Assessment.

Yoshinaga-Itano, C., Baca, R. L., & Sedey, A. L. (2010). Describing the trajectory of language development in the presence of severe-to-profound hearing loss: A closer look at children with cochlear implants versus hearing aids. *Otology & Neurotology*, *31*(8), 1268–1274. https://doi.org/10.1097/MAO.0b013e3181f1ce07

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2006). *Preschool Language Scale, 4th edition*. Pearson Education Inc.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scale, 5th edition*. Pearson Education Inc.

Appendix A

CINAHL Search Strategy

#1	(MH “Outcome Assessment”) OR (MH “Outcomes (Health Care)”) OR (MH “Treatment Outcomes”)
#2	(MH “Child, Disabled”) OR (MH “Child, Preschool”) OR (MH “Child Health”) OR (MH “Child Development Disorders”)
#3	(MH “Hearing Loss, Functional”) OR (MH “Hearing Loss, Partial”) OR (MH “Hearing Loss, Sensorineural”) OR (MH “Hearing Loss, Conductive”) OR (MH “Hearing Disorders”) OR (MH “Deafness”)
#4	(MH “Language”) OR (MH “Speech and Language Assessment”) OR (MH “Rehabilitation, Speech and Language”) OR (MH “Language Disorders”
#5	(MH “Outcome Assessment”) OR (MH “Outcomes (Health Care)”) OR (MH “Treatment Outcomes”)
#6	(MH “Child, Disabled”) OR (MH “Child, Preschool”) OR (MH “Child Health”) OR (MH “Child Development Disorders”)
#7	(MH “Hearing Loss, Functional”) OR (MH “Hearing Loss, Partial”) OR (MH “Hearing Loss, Sensorineural”) OR (MH “Hearing Loss, Conductive”) OR (MH “Hearing Disorders”) OR (MH “Deafness”)
#8	(MH “Language”) OR (MH “Speech and Language Assessment”) OR (MH “Rehabilitation, Speech and Language”) OR (MH “Language Disorders”)
#9	S5 AND S6 AND S7 AND S8
#10	(MH “Clinical Assessment Tools”) OR (MH “Speech and Language Assessment”) OR (MH “Outcome Assessment”) OR (MH “Functional Assessment”)
#11	(MH “Instrument Validation”)
#12	(MH “Clinical Assessment Tools”)
#13	(MH “Language Tests”)
#14	((MH “Language Tests”)) AND (S1 OR S10 OR S11 OR S12 OR S13)
#15	(((MH “Language Tests”)) AND (S1 OR S10 OR S11 OR S12 OR S13)) AND (S5 OR S14)

Appendix A (cont.)

#16 ((((MH “Language Tests”)) AND (S1 OR S10 OR S11 OR S12 OR S13)) AND (S5 OR S14)) AND (S6 AND S7 AND S8 AND S15)

#17 (MH “Measurement Issues and Assessments”)

#18 ((MH “Measurement Issues and Assessments”)) AND (S1 OR S5 OR S10 OR S11 OR S12 OR S13 OR S17)

#19 ((MH “Measurement Issues and Assessments” OR S1 OR S5 OR S10 OR S11 OR S12 OR S13 OR S17)

#20 (((MH “Measurement Issues and Assessments” OR S1 OR S5 OR S10 OR S11 OR S12 OR S13 OR S17)) AND (S3 AND S4 AND S6 AND S19)

#21 (((MH “Measurement Issues and Assessments” OR S1 OR S5 OR S10 OR S11 OR S12 OR S13 OR S17)) AND (S3 AND S4 AND S6 AND S19)

#22 (MH “Infant”) OR (MH “Infant Development”)

#23 (MH “Early Childhood Intervention”)

#24 ((MH “Early Childhood Intervention”)) OR (S6 OR S22 OR S23)

#25 (((MH “Early Childhood Intervention”)) OR (S6 OR S22 OR S23)) AND (S7 AND S8 AND S19 AND S24)

Appendix B

Supplemental materials describing the 36 studies can be found on Open Sciences Framework https://osf.io/ncm23/?view_only=1455217c19c44e3881e4628ed252fe3a

Details such as study authors, tests used, sample characteristics, and study purposes are laid out in an easy-to-read table. We also list whether the authors included composite scores, made group comparisons, noted informal differences, and evaluated change over time. Finally, we noted if the study had statistically significant or significant results or if they included other analyses.

Supplemental Materials References

Ambrose, S. E., Unflat Berry, L. M., Walker, E. A., Harrison, M., Oleson, J., & Moeller, M. P. (2014). Speech sound production in 2-year-olds who are hard of hearing. *American Journal of Speech-Language Pathology*, *23*(2), 91–104. https://doi.org/10.1044/2014_AJSLP-13-0039

Apuzzo, M., & Yoshinaga-Itano, C. (1995). Early identification of infants with significant hearing loss and the Minnesota Child Development Inventory. *Seminars in Hearing*, *16*(2), 124–135. https://doi.org/10.1055/s-0028-1083710

Bentler, R., Walker, E., McCreery, R., Arenas, R. M., & Roush, P. (2014). Nonlinear frequency compression in hearing aids: Impact on speech and language development. *Ear and Hearing*, *35*(4), e143–e152. https://doi.org/10.1097/AUD.0000000000000030

Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., Wright, M., Psarros, C., Rattigan, K., & 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*, 264–285. https://doi.org/10.1044/1092-4388(2001/022)

Ching, T. Y. C., Crowe, K., Martin, V., Day, J., Mahler, N., Youn, S., Street, L., Cook, C., & 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. https://doi.org/10.3109/17549500903577022

Ching, T. Y. C., Day, J., Dillon, H., Gardner-Berry, K., Hou, S., Seeto, M., Wong, A., & Zhang, V. (2013). Impact of the presence of auditory neuropathy spectrum disorder (ANSD) on outcomes of children at three years of age. *International Journal of Audiology*, *52*(sup2), S55–S64. https://doi.org/10.3109/14992027.2013.796532

Ching, T. Y. C., Day, J., Zhang, V., Dillon, H., Van Buynder, P., Seeto, M., Hou, S., Marnane, V., Thomson, J., Street, L., Wong, A., Burns, L., & Flynn, C. (2013). A randomized controlled trial of nonlinear frequency compression versus conventional processing in hearing aids: Speech and language of children at three years of age. *International Journal of Audiology*, *52*(sup2), S46–S54. https://doi.org/10.3109/14992027.2013.858281

Ching, T. Y. C., Dillon, H., Button, L., Seeto, M., Van Buynder, P., Marnane, V., Cupples, L., & Leigh, G. (2017). Age at intervention for permanent hearing loss and 5-year language outcomes. *Pediatrics*, *140*(3), e20164274. https://doi.org/10.1542/peds.2016-4274

Ching, T. Y. C., Dillon, H., Hou, S., Zhang, V., Day, J., Crowe, K., Marnane, V., Street, L., Burns, L., Van Buynder, P., Flynn, C., & Thomson, J. (2013). A randomized controlled comparison of NAL and DSL prescriptions for young children: Hearing-aid characteristics and performance outcomes at three years of age. *International Journal of Audiology*, *52*(sup2), S17–S28. http://doi.org/10.3109/14992027.2012.705903

Ching, T. Y. C., Zhang, V. W., Johnson, E. E., Van Buynder, P., Hou, S., Burns, L., Button, L., Flynn, C., & McGhie, K. (2018). Hearing aid fitting and developmental outcomes of children fit according to either the NAL or DSL prescription: Fit-to-target, audibility, speech and language abilities. *International Journal of Audiology*, *57*(sup2), S41–S54. https://doi.org/10.1080/14992027.2017.1380851

Cupples, L., Ching, T. Y. C., Button, L., Seeto, M., Zhang, V., Whitfield, J., Gunnourie, M., Martin, L., & Marnane, V. (2018). Spoken language and everyday functioning in 5-year-old children using hearing aids or cochlear implants. *International Journal of Audiology*, *57*(Supp. 2), S55–S69. https://doi.org/10.1080/14992027.2017.1370140

Cupples, L., Ching, T. Y. C., Crowe, K., Day, J., & Seeto, M. (2014). Predictors of early reading skill in 5-year-old children with hearing loss who use spoken language. *Reading Research Quarterly*, *49*(1), 85–104. https://doi.org/10.1002/rrq.60

Daub, O., Bagatto, M. P., Johnson, A. M., & Oram Cardy, J. (2017). Language outcomes in children who are deaf and hard of hearing: The role of language ability before hearing aid intervention. *Journal of Speech, Language, and Hearing Research*, *60*(11), 3310–3320. https://doi.org/10.1044/2017_JSLHR-L-16-0222

DesJardin, J. L., Doll, E. R., Stika, C. J., Eisenberg, L. S., Johnson, K. J., Ganguly, D. H., Colson, B. G., & Henning, S. C. (2014). Parental support for language development during joint book reading for young

children with hearing loss. *Communication Disorders Quarterly*, *35*(3), 167–181. https://doi.org/10.1177/1525740113518062

Eriks-Brophy, A., Gibson, S., & Tucker, S.-K. (2013). Articulatory error patterns and phonological process use of preschool children with and without hearing loss. *Volta Review, 113*(2), 87–125.

Fitzpatrick, E. M., Crawford, L., Ni, A., & Durieux-Smith, A. (2011). A descriptive analysis of language and speech skills in 4- to 5-year-old children with hearing loss. *Ear and Hearing*, *32*(5), 605–616. https://doi.org/10.1097/AUD.0b013e31821348ae

Fitzpatrick, E. M., Durieux-Smith, A., Gaboury, I., Coyle, D., & Whittingham, J. (2015). Communication development in early-identified children with mild bilateral and unilateral hearing loss. *American Journal of Audiology*, *24*(3), 349–353. https://doi.org/10.1044/2015_AJA-15-0003

Fitzpatrick, E. M., Gaboury, I., Durieux-Smith, A., Coyle, D., Whittingham, J., & Nassrallah, F. (2019). Auditory and language outcomes in children with unilateral hearing loss. *Hearing Research*, *372*, 42–51. https://doi.org/10.1016/j.heares.2018.03.015

Hogan, S., Stokes, J., White, C., Tyszkiewicz, E., & Woolgar, A. (2008). An evaluation of Auditory Verbal therapy using the rate of early language development as an outcome measure. *Deafness & Education International*, *10*(3), 143–167. https://doi.org/10.1002/dei.242

Jackson, C. W., & Schatschneider, C. (2014). Rate of language growth in children with hearing loss in an auditory-verbal early intervention program. *American Annals of the Deaf*, *158*(5), 539–554. https://doi.org/10.1353/aad.2014.0006

Koehlinger, K., Van Horne, A. O., Oleson, J., McCreery, R., & Moeller, M. P. (2015). The role of sentence position, allomorph, and morpheme type on accurate use of s-related morphemes by children who are hard of hearing. *Journal of Speech, Language, and Hearing Research*, *58*(2), 396–409. https://doi.org/10.1044/2015_JSLHR-L-14-0134

Marciano, E., Laria, C., Malesci, R., Iadicicco, E., Landolfi, E., Niri, C., Papa, C., Franze, A., Auletta, G. (2013). Newborn hearing screening in the Campania region (Italy): Early language and perceptual outcomes of infants with permanent hearing loss. *Acta Otorhinolaryngologica Italica, 33,* 414–417.

Mayne, A. M., Yoshinaga-Itano, C., & Sedey, A. L. (1999). Receptive vocabulary development of infants and toddlers who are deaf or hard of hearing. *The Volta Review, 100,* 29–52.

Mayne, A. M., Yoshinaga-Itano, C., Sedey, A. L., & Carey, A. (1998). Expressive vocabulary development of infants and toddlers who are deaf or hard of hearing. *The Volta Review, 100*, 1–28.

McCreery, R. W., Walker, E. A., Spratford, M., Oleson, J., Bentler, R., Holte, L., & Roush, P. (2015). Speech recognition and parent ratings from auditory development questionnaires in children who are hard of hearing: *Ear and Hearing*, *36*, 60S–75S. https://doi.org/10.1097/AUD.0000000000000213

Moeller, M. P. (2000). Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics*, *106*(3), e43–e51. https://doi.org/10.1542/peds.106.3.e43

Pipp-Siegel, S., Sedey, A. L., VanLeeuwen, A. M., & Yoshinaga-Itano, C. (2003). Mastery motivation and expressive language in young children with hearing loss. *Journal of Deaf Studies and Deaf Education*, *8*(2), 133–145. https://doi.org/10.1093/deafed/eng008

Pressman, L., Pipp-Siegal, S., Yoshinaga-Itano, C., & Deas, A. (1999). Maternal sensitivity predicts language gain in preschool children who are deaf and hard of hearing. *Journal of Deaf Studies and Deaf Education*, *4*(4), 294–304. https://doi.org/10.1093/deafed/4.4.294

Pressman, L. J., Pipp-Siegal, S., Yoshinaga-Itano, C., Kubicek, L., Emde, R. N. (1999). A comparison of the links between emotional availability and language gain in young children with and without hearing loss. *Volta Review, 100*(5), 251–277.

Tomblin, J. B., Harrison, M., Ambrose, S. E., Walker, E. A., Oleson, J. J., & Moeller, M. P. (2015). Language outcomes in young children with mild to severe hearing loss: *Ear and Hearing*, *36*, 76S–91S. https://doi.org/10.1097/AUD.0000000000000219

Tomblin, J. B., Oleson, J. J., Ambrose, S. E., Walker, E., & Moeller, M. P. (2014). The influence of hearing aids on the speech and language development of children with hearing loss. *JAMA Otolaryngology–Head & Neck Surgery*, *140*(5), 403. https://doi.org/10.1001/jamaoto.2014.267

Vohr, B., Jodoin-Krauzyk, J., Tucker, R., Topol, D., Johnson, M. J., Ahlgren, M., & Pierre, LSt. (2011). Expressive vocabulary of children with hearing loss in the first 2 years of life: Impact of early intervention. *Journal of Perinatology*, *31*(4), 274–280. https://doi.org/10.1038/jp.2010.110

Walker, E., McCreery, R., Spratford, M., & Roush, P. (2016). Children with auditory neuropathy spectrum disorder fitted with hearing aids applying the American Academy of Audiology pediatric amplification guideline: Current practice and outcomes. *Journal of the American Academy of Audiology*, *27*(3), 204–218. https://doi.org/10.3766/jaaa.15050

Yoshinaga-Itano, C., Baca, R. L., & Sedey, A. L. (2010). Describing the trajectory of language development in the presence of severe-to-profound hearing loss: A closer look at children with cochlear implants versus hearing aids. *Otology & Neurotology*, *31*(8), 1268–1274. https://doi.org/10.1097/MAO.0b013e3181f1ce07

Yoshinaga-Itano, C., & Sedey, A. (1998). Early speech development in children who are deaf or hard of hearing: Interrelationships with language and hearing. *Volta Review, 100*(5), 181–211.

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. https://doi.org/10.1542/peds.102.5.1161

EHDInfo [Can be downloaded here for distribution.]

It may bring comfort to know what specific things your provider (audiologist, health care clinician, early intervention specialist, etc.) is doing to keep you and your baby safe. Your provider may also ask you to take certain steps to keep them and their staff safe. Many providers are calling families prior to their appointment to discuss safety.

Emerging Solutions: How to Keep You, Your Baby, and Your Provider Safe During COVID-19

Staying Safe During Your Appointment

During the call with your provider, consider asking:

1. If doing a hearing screening only, do you have screening options other than us entering the building (e.g., screening in car)?
2. If there is paperwork to be filled out, can you send it to me ahead of time?
3. When I arrive, are there specific instructions (e.g., phone before I enter the building)?
4. Is there a limit to who can come to the appointment with me and my child?
5. Is there a limit to the number of people who can be in the waiting area?
6. Are there health screenings (e.g., temperature) of patients upon arrival?
7. How are public areas being cleaned (e.g., waiting rooms, restrooms, food service areas) and how often?
8. How do you screen yourself or staff for wellness (e.g., temperature)?
9. What protective gear (e.g., gloves, masks) does the provider and his/her staff use?
10. How is equipment (e.g., screening, diagnostic) cleaned or replaced between patients?
11. If you will be talking directly to my child, do you have a face mask with clear plastic so that my child can see your face/lips?
12. How can I help keep you and your staff safe?
    - Would you like me to wear a face mask?
    - If the clinic serves both sick and well patients, how will you handle that?
    - Anything else?

If You Decide to Cancel or Reschedule

Even though your provider is taking steps toward safety, if you still do not feel comfortable with an in-person appointment, you may want to think about and/or take action in the following ways:

1. Have you talked to your provider about:
    - Your safety concerns?
    - Additional safety strategies that would make you more comfortable to attend an appointment?
2. Would it help to talk to another parent who has recently had the experience of an in-person appointment?
3. If you plan to cancel or reschedule, and you have an appointment scheduled, please call and let your provider know at least 48 hours in advance (or within the timeframe outlined by your provider). Not showing up impacts the schedule of the provider and his/her staff.
4. If you plan to reschedule your appointment:
    - Ask your provider how far out they are scheduled.
    - Have you balanced your concerns with safety with the amount of time that will pass until you are able to be seen by your provider?
    - Does the delay in going to the appointment impact the services your child needs?
5. Ask your provider if they can do a video visit by a secured system.

> We went to the audiologist at our CI Center last week, and I've been ***VERY*** anxious about COVID. It was a ***VERY*** comfortable experience!!! The CI Center called us when they were ready to re-open. They were very transparent about the new policies (masks, temp checks, etc.) and wanted me to know that I could cancel at the last minute if I wasn't comfortable. There was no waiting room—only waiting in the vehicle was allowed. There were cones lined up in the parking lot with phone numbers and spot numbers on them. You let them know what spot you were parked at, and they came out, with PPE on, with extra masks if we didn't have any. They took our temperatures and asked us some questions. They gave us hand sanitizer, and we went into the appointment. ***LOTS*** of sanitizer was used by the audiologists, and everything that was touched was thrown away or set aside for sterilization. We didn't need to check out. Everything was done over the phone after the appointment. It was a ***LOVELY*** experience for this COVID-anxious mama!
>
> *—Michelle Thomas, Parent, Michigan*

Additional Resources

- https://www.healthyhearing.com/report/53087-Need-to-see-a-hearing-care-specialist-during-the-pandemic-things-to-keep-in-mind
- https://www.asha.org/News/2020/Early-Intervention-and-COVID-19-Advice-for-Parents-of-Children-Ages-0–3-Whose-Services-Are-Interrupted/
- https://handsandvoices.org/fl3/topics/fam-fam-support/need-support.html
- http://www.handsandvoices.org/fl3/topics/tipsheets.html
- http://www.infanthearing.org/COVID-19/index.html#support

The development and distribution of this material was supported in part by the Maternal and Child Health Bureau (MCHB) of the Health Resources and Services Administration (HRSA) as part of award U52MC0439, totaling $3,400,000; and as part of award 2UJ1MC307480400, totaling $1,800,000. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement by, HRSA, HHS, or the US Government.