Multidisciplinary assessment and interventions for childhood auditory processing disorder (APD) and listening difficulties (LiD)

Authors

David R. Moore,a, b, c
Li Lin,a
Ritu Bhalerao,a
Jody Caldwell-Kurtzman,a
Lisa L. Hunter,a, b

Affiliations

aCommunication Sciences Research Center, Division of Patient Services Research, Cincinnati Children’s Hospital Medical Center, OH
bDepartment of Otolaryngology-Head & Neck Surgery, University of Cincinnati, OH
cManchester Centre for Audiology and Deafness, University of Manchester, United Kingdom

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Purpose: Auditory processing disorder (APD) has been studied in both research and clinic settings, but the relation between the two has not been addressed. In a longitudinal research study (SICLiD), we found that children with clinically normal audiometry who had caregiver-reported listening difficulties (LiD), with or without clinically assessed APD, performed poorly on both listening and cognitive tests. Specific questions asked here were, for the children with LiD, what other neurodevelopmental clinical conditions were identified, what interventions were used by different clinical providers, and how clinical practice was predicted by research results.

Methods: Study setting was a large, research-led, tertiary pediatric hospital. Electronic medical records of 74 children aged 6-13 years, recruited into SICLiD and assigned to an LiD group based on a validated and reliable caregiver report (ECLiPS), were independently reviewed. Focus was on clinical assessments and interventions following appointments provided in the Hospital Divisions of Audiology, Occupational Therapy, Psychology (Developmental and Behavioral Pediatrics), and Speech-Language Pathology (SLP), prior to participation in SICLiD. Descriptive statistics on clinical encounters, identified conditions, and interventions were compared with quantitative, standardized performance on SICLiD assessments of listening and cognitive function. SICLiD z-scores were compared for participants with and without each clinical condition using univariate and logistic prediction analyses.

Results: Most (86%) of the children with LiD had been evaluated by at least one clinical service. Overall, 24 assessment categories related to LiD, including APD, were identified. Most common conditions were attention (32%), language (28%), hearing (18%), anxiety (16%), and autism spectrum (6%) disorders. Performance on SICLiD measures varied significantly between providers, conditions, and interventions. Significant relationships between SICLiD and clinical conditions were mostly caregiver-reported items from the ECLiPS or the Children's Communication Checklist (CCC-2). Other significant correlations were scarce, but included the
SCAN composite score, which predicted clinical language and attention, but not other auditory abilities or APD. SICLiD data combined with caregiver reports provided reliable predictions of all clinical conditions except APD.

Conclusions: The variety of disciplines, assessments, conditions and interventions revealed here supports previous studies showing that LiD and APD are multifaceted problems of neurodevelopment. Comparisons between clinical- and research-based assessments suggest a diagnostic path that prioritizes caregiver reports and selected psychometric tests for screening and diagnostic purposes.
Abbreviations

APD – Auditory Processing Disorder
AUROC - Area under the receiver operating characteristic
CCC-2 - Children’s Communication Checklist (2nd edition)
CCC GCC - Children’s Communication Checklist – General Communication Composite
CCC SIDI - Children’s Communication Checklist – Social interaction difference index
CCHMC – Cincinnati Children’s Hospital Medical Center
DL - Bergen Dichotic Listening Test
DL I -15 – DL laterality score with an ILD = -15 dB
ECLiPS EAS - ECLiPS Environmental and Auditory Sensitivity subscale
ECLiPS – Everyday Children’s Listening and Processing Scale
ECLiPS LLL - ECLiPS Literacy/Language/Laterality subscale
ECLiPS MA - ECLiPS Memory and Attention subscale
ECLiPS PSS – ECLiPS Pragmatic and Social Skills subscale
ECLiPS SAP – ECLiPS Speech and Auditory Processing subscale
NIH FC – Fluid reasoning Composite score of NIH Cognition Toolbox
NIH FT – Flanker test of NIH Cognition Toolbox
ILD – Interaural level difference
LiD – Listening Difficulties
LiSN-S – Listening in Spatialized Noise – Sentences test
LiSN HC - LiSN-S High Cue subscale
<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>71</td>
<td>LiSN LC – LiSN-S Low Cue subscale</td>
</tr>
<tr>
<td>72</td>
<td>LiSN SA - LiSN-S Spatial Advantage subscale</td>
</tr>
<tr>
<td>73</td>
<td>LiSN TA - LiSN-S Talker Advantage subscale</td>
</tr>
<tr>
<td>74</td>
<td>NIH - National Institutes of Health Cognition Toolbox</td>
</tr>
<tr>
<td>75</td>
<td>NIH CC – Crystalized reasoning Composite score of NIH Cognition Toolbox</td>
</tr>
<tr>
<td>76</td>
<td>NIH DCCS – Directional Change Card Sorting test of NIH Cognition Toolbox</td>
</tr>
<tr>
<td>77</td>
<td>NIH ECC – Early Childhood Composite score of NIH Cognition Toolbox</td>
</tr>
<tr>
<td>78</td>
<td>NIH LS – List sorting test of NIH Cognition Toolbox</td>
</tr>
<tr>
<td>79</td>
<td>NIH ORR – Oral reading recognition test of NIH Cognition Toolbox</td>
</tr>
<tr>
<td>80</td>
<td>NIH PC – Pattern comparison processing speed test of NIH Cognition Toolbox</td>
</tr>
<tr>
<td>81</td>
<td>NIH PV – Picture Vocabulary test of NIH Cognition Toolbox</td>
</tr>
<tr>
<td>82</td>
<td>NIH TC – Total Composite score of NIH Cognition Toolbox</td>
</tr>
<tr>
<td>83</td>
<td>OR – Odds ratio</td>
</tr>
<tr>
<td>84</td>
<td>PE – Tympanic pressure equalization tubes</td>
</tr>
<tr>
<td>85</td>
<td>SCAN AFG - SCAN Auditory Figure-Ground subscale</td>
</tr>
<tr>
<td>86</td>
<td>SCAN FW – SCAN Filtered Words subscale</td>
</tr>
<tr>
<td>87</td>
<td>SCAN CS - SCAN Competing Sentences subscale</td>
</tr>
<tr>
<td>88</td>
<td>SCAN CWD - SCAN Competing Words subscale</td>
</tr>
<tr>
<td>89</td>
<td>SICLiD – Sensitive Indicators of Childhood Listening Difficulties (study)</td>
</tr>
<tr>
<td>90</td>
<td>TD – Typically Developing</td>
</tr>
<tr>
<td>91</td>
<td>VIF - Variance inflation factor</td>
</tr>
<tr>
<td>92</td>
<td></td>
</tr>
</tbody>
</table>
Introduction

Auditory processing disorder (APD) is a deficit of active listening and is believed to originate in the brain, not the ear (AAA, 2010; ASHA, 2005), since people diagnosed with APD typically do not have a clinical hearing loss or otological disorders (Hunter et al., 2021, 2023; Moore et al., 2013; Petley et al., 2024). There have been many debates and discussions among health professionals and researchers about APD diagnosis (Dillon & Cameron, 2021; Illiadou & Kiese-Himmel, 2018; Moore, 2018) and treatment (Bellis et al., 2012; Fey et al., 2011) in school-aged children. APD overlaps with developmental language disorders, attention deficit hyperactivity disorders, autism spectrum disorders, and other problems (Moore et al., 2018), to the extent that children diagnosed with APD do not, on average, perform differently on many listening, language and cognitive tests from children assessed to have other neurodevelopmental conditions (Dawes & Bishop, 2009; de Wit et al., 2018; Ferguson et al., 2011; Sharma et al., 2009). A long-standing issue is that, due to lack of a solid theoretical basis, established physiological mechanisms, or consensus on the definition of APD, appropriate diagnostic testing, and intervention cannot be rationalized (Papesh et al., 2023; Vermiglio, 2014). An alternative or umbrella term for APD, “listening difficulties” (LiD), focuses on symptoms commonly reported by children and their caregivers or adults who present at audiology clinics, but have clinically normal audiograms (Barry & Moore, 2021; Dillon & Cameron, 2021; Petley et al., 2021; Shiels et al., 2023). This term, LiD, recognizes that the underlying causes are multifactorial and not necessarily due to auditory system dysfunction.

Interventions for LiD in children include wireless, remote microphone hearing devices (Johnston et al., 2009; Lemos, 2009; Shiels et al., 2023), aural rehabilitation (Boothroyd, 2007), usually in the form of computer game-based training (e.g. Fast ForWord; Strong et al., 2011).

---

1 In this paper we use the term “condition” to refer to all clinical statements of impaired function since it was often unclear whether a formal diagnosis had been made. The exception is APD for which a clinical diagnosis was made in some cases. The term “assessment” is used to refer to research or clinical tests.
“metacognitive training” (Chermak & Musiek, 2013), classroom accommodations (DeBonis, 2015), and other listening exercises including dichotic listening, language and sound discrimination training (Fey et al., 2011; Lotfi Y, 2016; Moncrieff et al., 2017; Sharma et al., 2012). The level of evidence for the effectiveness of these interventions is variable but, in general, does not support their widespread adoption (Fey et al., 2011). One series of studies found, however, that training using a spatialized sound task (LiSN & Learn) resulted in improved spatial hearing for children identified with ‘spatial processing disorder’ (Cameron & Dillon, 2011). A gold-standard, blinded control method, and independently developed self- and caregiver-report questionnaires, suggested in subsequent studies that more generalized aspects of listening may also be improved (Cameron et al., 2012; Cameron et al., 2015). Recently, Shiels et al. (2023) provided strong evidence for the effectiveness of remote microphone devices in aiding LiD. To date, however, no intervention method has seen general acceptance. High quality, large-scale clinical trials have not been performed and, as for assessment, there is no consensus on what outcome measures to use. Consequently, LiD is a challenging condition for parents, providers, and insurers to understand, evaluate, endorse, and fund.

To address questions about the underlying mechanisms of childhood LiD, Cincinnati Children’s Research Foundation (CCRF) and the National Institute of Deafness and Communication Disorders (NIDCD), funded a longitudinal project, Sensitive Indicators of Childhood Listening Difficulties (SICLiD). Children (6-13 years old) with clinically normal audiometry were recruited into two groups, listening difficulty (LiD) or typically developing (TD), based primarily on the results of a validated and reliable caregiver evaluation instrument (the ECLiPS; Barry & Moore, 2021; Denys et al., 2024; Petley et al., 2021). A key premise of the ECLiPS is that caregivers, typically parents, observe their children’s behavior over extended periods in everyday life and

2 now known as “Sound Storm” (https://www.soundstorm.app/)
are therefore best able to judge the child’s listening and related abilities. A similar rationale was used to develop a widely-used caregiver report of children’s language abilities (CCC-2; Bishop, 2006) that inspired the ECLiPS design. The ECLiPS has 38 statements (items) that various stakeholder groups suggested are characteristics of children with LiD and that the caregiver rates on a 5-point Likert scale from ‘strongly agree’ to ‘strongly disagree’. Children assigned to both TD and LiD groups were tested on a range of audiological and other physiological and behavioral measures. Papers published to date have focused on the results of those tests (Hunter et al., 2021, 2023; Moore et al., 2020; Petley et al., 2021, 2024; Stewart et al., 2022).

They suggest that LiD/APD is a result primarily of generally impaired cognitive processes (e.g. attention, memory, executive function), and speech-language processing mechanisms beyond the central auditory nervous system.

The families of children with LiD who participated in the SICLiD project were approached to consent for their child’s relevant medical history to be accessed by this study’s investigators. The overall goal of that access, and of the study reported here, was to investigate the hearing and related assessments and interventions the children have received within the clinical divisions of Cincinnati Children’s Hospital Medical Center (CCHMC). Nearly all the clinical assessments and interventions had been provided before the children entered the SICLiD study and, in all cases, the SICLiD results were unknown to the clinical providers. We hypothesized, based on previous studies, that the children would have received a wide variety of assessments, but relatively few and diverse interventions. Our specific questions were: What type and how many clinical assessments were given? How do those clinical assessments relate to the results of the SICLiD study? What type of interventions have been used?

Methods

Participants and caregiver reports
Participants were 74 children, aged 6-13 years at baseline assessment, and previously determined in the SICLiD research study to have significant LiD as defined by the Total ECLiPS score (Petley et al., 2021). Caregivers of all recruited children also completed the CCC-2 (Bishop, 2006), and a background questionnaire concerning demographics (age, race/ethnicity, caregiver education level), history of neurologic and otologic appointments, and diagnoses or assessments of learning disorders. Those with severe neurologic, otologic, or intellectual conditions (e.g. current use of tympanic pressure equalizing (PE) tubes, unable to complete tasks) were excluded, but those with common neurodevelopmental disorders known to overlap with LiD (e.g. language, attention, executive function, high-functioning autism spectrum disorder - ASD, history of otitis media) were included.

**Hearing and cognitive tests**

Both clinical testing and research were conducted within CCHMC. All 74 children had an extensive audiometric assessment (Hunter et al., 2021), including pure tone audiograms (octave frequencies from 0.25 – 8.0 kHz, plus 10, 12.5, and 16 kHz, bilaterally), wideband tympanometry, acoustic reflex growth, transient and distortion product otoacoustic emissions, as well as auditory brainstem and cortical evoked responses (Hunter et al., 2023, Petley et al., 2023). In other behavioral assessments, standardized scores on the SCAN-3:C (Keith, 2009), the test suite most commonly used in the US to diagnose APD (Emanuel, Ficca, & Korczak, 2011), and on all 4 SCAN subtests, were obtained. Standardized scores on the Low- and High-Cue, and the Spatial, Talker and Total Advantage subtests of the Listening in Spatialized Noise: Sentences Test (LiSN-S; Cameron & Dillon, 2007), and on all subtests and Composites of the National Institutes of Health (NIH) Cognition Toolbox (Weintraub et al., 2013) were also obtained (see Petley et al., 2021, for further details). Performance on the Bergen Dichotic Listening Test (DL; Hugdahl et al., 2009; Moore et al., 2020) was available for a sub-set (n=44) of the children.
Hospital records

For the 64/74 children with LiD who had an electronic medical record (Epic) that included assessments of hearing, listening, and/or other developmental disorders, a retrospective analysis of each child’s record was completed. Each service operated differently with respect to the clinical assessment. The diagnostic result was often not clearly documented, and many different terms were used to describe the child’s identified problems and intervention recommendations. For this reason, we are using the term “condition” rather than “diagnosis”, to categorize the problems identified for each participant. For each participant, a summary description of conditions related to APD/LiD was first noted. Then, for each clinical Division - Speech-language pathology (SLP), Audiology, Developmental and Behavioral Pediatrics (Psychology\(^3\)), Physical and Occupational Therapy (OT), and for each participant, we extracted and coded the following data: appointment type (assessment, therapy), goals, number of encounters, progress/outcome. Assessment and therapy types were coded (up to 6 per Division for each child). Multiple authors independently performed repeat summaries and counts of the EPIC records, overseen by senior authors (LLH, DRM) who made the final decision on how to categorize each child’s records.

Analysis: Caregiver, hearing, listening, and cognitive scores were presented as charts or violin plots. Violin plots showed probability density of the data, and median, interquartile range and total range in the context of standardized scores expected from typically developing children.

Epic assessment data were manually coded and reduced to five summary conditions: Language (including speech, vocabulary, reading, apraxia/dyspraxia, stuttering), Hearing (including

\(^3\) The Department of Developmental and Behavioral Pediatrics at CCHMC covers a range of disciplines of which only the Psychology Group is relevant here. We therefore refer to this team as “Psychology” through the rest of the study.
APD/LiD, sensory processing⁴, hearing loss⁵), Attention (including attention deficit hyperactivity disorder, ADHD), Anxiety (including panic, stress, post-traumatic stress disorder, PTSD, and depression), and ASD (including social or pragmatic problems). Each child was assigned a binary code (0 – no symptoms, 1 – symptoms; Table 1) for each condition.

All SICLiD scores were converted to age-specific Z-scores so that the relation between each score was equalized to ascertain prediction level for each clinical condition. Mean Z-scores were first compared between children without or with symptoms of each condition (0 or 1 above) using two sample or Satterthwaite t-tests to assess the univariate statistical significance (p-value) and effect size (d) of the difference (Table 1). Critical p-values were assessed using Benjamini-Hochberg adjustments to determine a false discovery rate < 5%. However, because of the relatively small number of SICLiD tests that were predictors, additional tests with marginal statistical significance (0.05 < p < 0.1) were also included in Table 1. Data of children assessed for APD were, additionally, analyzed separately using multiple t tests comparing scores on all SICLiD measures between children with (n = 22) and without (n = 22) an APD diagnosis (including auditory processing ‘weakness’; see Moore et al., 2018).

SICLiD univariate predictors listed in Table 1 were considered in a subsequent adjusted logistic regression model of clinical assessments. Relationships and multicollinearity among the predictors were explored using the correlation coefficient and variance inflation factor (VIF). A high correlation and a VIF greater than 5 indicates problematic multicollinearity. The final logistic prediction models (Table 2) were selected with the best prediction combinations of SICLiD scores, reported with odds ratios (with 95% confidence intervals) and an area under the receiver

---

⁴ “Sensory processing” and “sensory processing disorder” are typically assessed by occupational therapists. They may involve any or all of the senses and can include up to 6 subtypes (STAR Institute). We analyzed the “Hearing and Auditory Processing” assessment with and without the 8 children who received this assessment.

⁵ All of the children in this study were found in the SICLiD study to have clinically normal hearing. However, 3 of them had been previously assessed with conductive hearing loss.
operating characteristic (AUROC) curve. A recent review of the clinical AUROC literature warned against labeling results (e.g. ‘good discrimination’; de Hond et al., 2022), so only AUROC values are presented here. Data analysis was conducted using SAS statistical software, version 9.4 (SAS Institute, Cary, N.C.).

**Results**

*Research tests of hearing, listening and cognition*

Performance on SICLiD tests of hearing, listening, and cognitive function for this sample of children with LiD is summarized in Fig. 1. Performance of the entire SICLiD sample (n = 146) on behavioral tests is reported in detail elsewhere (Hunter et al., 2021; Kojima, 2024; Petley et al., 2021). All participants with LiD had clinically normal hearing, including pure tone audiograms (PTA ≤ 20 dB HL bilaterally at each octave frequency from 0.25 – 8.0 kHz). Additional, extended high frequency testing (10, 12.5, and 16 kHz, bilaterally) revealed substantial variability, with some hyper-acute (-15 dB HL) and insensitive (to 50 dB HL) thresholds (Fig. 1A). Total ECLiPS scores were uniformly low, since this was the primary inclusion criterion. ECLiPS scales (SAP etc.) were more variable than the Total, but all had a median score between 3-5, well below the median standard score of 10 (Fig. 1B, D). CCC-2 “general communication composite” (GCC) scores correlated with Total ECLiPS scores (r = 0.39; p < 0.001) and revealed associated language problems of these children (Figs. 1B,C). CCC-SIDI scores (Fig. 1C) were above the typical range (> +10) for 23/74 participants, suggestive of a pragmatic language impairment, and below typical (< +10) for 9/74 participants, suggestive of ASD (Bishop, 2003). The SCAN suite of tests showed generally poor performance on two tests of dichotic listening, the Competing Words and Competing Sentences, average performance (10) on a speech-in-noise test (Auditory Figure-Ground) and above average performance on a test of low-pass Filtered Words (Fig. 1E). The Listening in Spatialized Noise – Sentences (LiSN-S) showed that children with LiD generally performed poorly in this complex, competing speech test. This was the case
for both individual tests (low cue, high cue), and for the Spatial Advantage and Talker Advantage derived measures, where cognitive aspects of performance are minimized (Fig. 1F; see (Petley et al., 2021)), but not for the Total Advantage derived measure. A suite of cognitive tests, the NIH Cognition Toolbox (Weintraub et al., 2013), showed widespread and relatively uniform difficulty for performance on an array of mostly visual-based cognitive tasks (Fig. 1G). Performance on the various conditions of the Bergen DL (Moore et al., 2020) did not show any difference between groups, with one exception, in the condition (ILD = -15 dB) for which laterality was more pronounced in the LiD group. Performance differences between children with and without diagnosed APD on all SICLiD tests (n = 28) were non-significant (p > 0.05).

**Clinical encounters and assessments**

Number of encounters and assessments for each Hospital service is shown in Fig. 2. Ten of the children did not have a clinical record at CCHMC, although the caregiver background questionnaires of at least three suggested they had been assessed elsewhere with an attention or speech/language problem. Of the remaining 64, the mean number of services used by each participant was 2.0. 75% attended at least one appointment with Audiology, and (69%) with SLP, while 28% were seen by Psychology and 30% by OT (Fig. 2A). A total of 130 assessments of problem conditions of Attention, Language, Hearing (& ‘Processing’), Anxiety, and ASD were made among the sample of 64 children, a mean of 2.03 conditions per child. The most common condition was Attention, followed by Language, Hearing, Anxiety and high-functioning ASD. Two children had been assessed with all 5 conditions coded here (Fig. 2E). Aside from an overall Hearing condition, of whom 88% also had APD, the balance of Language, Attention, and ASD between the two APD groups was near identical. However, more than twice as many children without APD had an Anxiety assessment as those with APD (Figs. 2G,H). This difference was reduced when the three children with sensory processing disorder were removed (Fig. 2F). Anxiety is not considered further here as we did not have any SICLiD tests for it.
Clinical Intervention and outcomes

Although there is no consensus in the literature on what treatment children with LiD should be offered, we document here the follow-up that occurred in this sample. Interventions were highly heterogenous, and not consistently documented in the clinical records. The number of different interventions in SLP clinics (n = 36) was about twice that in any of the other three clinics (Fig. 2D). Some services recommended interventions to most of the children they saw (Psychology – 94%, SLP – 82%, OT – 68%), while Audiology recommended interventions to fewer than half the children they saw (48%). Audiology recommended interventions were most commonly behavioral (in order: accommodations, computer training, communication strategies), but devices, specifically remote microphone and/or low-gain hearing aids, were also popular. For the other services, a broad range of exclusively behavioral interventions was used. Most strikingly, both the average (Fig. 2B) and maximum (Fig. 2C) number of encounter sessions varied hugely between the specialties, with SLP (mean = 25) and OT (max = 187) having many intervention sessions, Psychology near the middle, and Audiology (mean = 3) having a small number of encounters that were mostly assessment. It was difficult to separate the intervention from the assessment, as these were often not specified separately. This challenge was particularly apparent in SLP, where the assessment and intervention often seemed to be one and the same thing. For example, assessment: “speech disturbance, articulation (/s/, /f/, /sh/, /z/), voice disturbance”, intervention: “work on speech, articulation and voice disturbances.”

It was often unclear whether the children participated in the stated interventions, or whether the interventions were only suggestions of procedures that could be tried. All the services except Audiology provided behavioral interventions during the encounter, whereas Audiology tended not to intervene for children with LiD but with clinically normal audiograms.

Scant evidence was available on outcomes, with almost no evidence of follow-up. However, SLP had the most consistent documentation on cessation of therapy. That service also had by
far the largest number of different specified interventions (n = 35 among 44 children), and the highest mean number of intervention sessions (24.9).

Relation between clinical assessments and research test scores

Following the methods described above, the prediction value for each clinical condition by each SICLiD research test was assessed. Of 80 univariate comparisons, only 16 showed significant differences (adjusted α < 0.05) between research scores based on binary clinical assessment category (i.e. without or with symptoms) and an effect size ≥ 0.5. An additional 5 comparisons with marginal statistical significance (0.1 > p > 0.05), two of which had p ≥ 0.5, are included in Table 1. Of these 21 comparisons, 9 were from the caregiver report measures (ECLiPS, CCC-2), 7 were from the SCAN, and 3 were from the NIH Oral reading test. Among conditions, Language, Attention, and Autism Spectrum had the largest numbers of predictive test scores, whereas the SCAN Composite score had the highest level of predictability (d = 0.92) of any single test, for Language.

Overall, clinical conditions were largely independent of research test scores, with the most predictive research scores for clinical groups coming from the language-centered caregiver report, the CCC-2. The CCC-2 also predicted clinical Language problems well . . . but not as well as the SCAN Composite, designed as a test for APD. For clinical Hearing and Processing, on the other hand, the caregiver ECLiPS SAP and CCC GCC were both moderately predictive, as expected. But none of the SCAN tests, nor any of the other SICLiD measures of speech-in-noise, cognition, or dichotic listening were significantly predictive of Hearing and Processing. The caregiver ECLiPS PSS and CCC SIDI, both reporting aspects of pragmatic language, also predicted clinical Attention and ASD-like symptoms, with the SCAN Composite and Competing Sentences (SCAN CS subtest) also predicting Attention. The NIH Oral Reading Recognition (ORR) and Picture Vocabulary (PV) Tests, designed primarily for measuring aspects of
Language, and the TAIL Distraction Test, designed for measuring Attention, were all predictive of ASD symptoms.

Children previously diagnosed (n=22) or not diagnosed (n=22) with APD did not show any significant or near-significant differences on performance of any of the research tasks. Consequently, no significant APD research predictors were identified.

**Predictive assessment logistic models**

Combining two or more research tests might provide better prediction of a clinical condition than just a single test. In this final analysis we created logistic regression models for each clinical condition in which at least marginal significant research predictors had been identified in the univariate analysis (Table 1). In all cases, the final models provided higher predictability than individual tests.

**Language**

The final prediction model consisted of the SCAN Composite and CCC SIDI (Table 2). SCAN Composite subscales (FW, AFG, CWD, CS) had high multicollinearity, with VIFs > 30 (see Methods). CCC GCC and LiSN LC subscale were highly correlated with the SCAN Composite (both p < 0.001). SCAN and LiSN subscales, and CCC GCC were therefore removed from the final model.

**Hearing and Processing**

The final model included the same predictors identified in the univariate analysis - CCC GCC and ECLiPS SAP (Table 2). Since the correlation between these scaled predictors of Hearing and Processing was non-significant, collinearity was not an issue.

**Attention**
The SCAN Comp was, again, a strong predictor of attention and achieved a relatively high AUROC when paired with the ECLiPS PSS (Table 2). Other potential pairings between these predictors and the ECLiPS EAS, the DLI15N and the CCC SIDI were rejected because of lower effect size, small sample size, and wide confidence intervals. The ECLiPS EAS also correlated highly with the ECLiPS PSS, although none of the predictors had VIF > 3.

ASD

The final prediction model included NIH OR and ECLiPS PSS, yielding a high AUROC (Table 2). No correlations between ASD predictors exceeded 0.5 and all had VIF < 1.5. Other combinations, including the TAIL Distract, also yielded impressive ASD prediction, but the TAIL sample size was small.

Discussion

Nearly all children with LiD recruited into the SICLiD study had previous encounters with clinical specialists working in audiology and related professions; 86% had attended relevant services at CCHMC, and caregiver reports for the remainder reported other professional interactions. As found in other, independent samples, and hypothesized here, most children had been seen by more than one type of clinical specialist due to multiple behavioral problems observed by caregivers (Ferguson et al., 2011; D.R. Moore et al., 2018; Sharma et al., 2009). Families of children with LiD thus have a wide range of concerns that extend well beyond the audiological domain. Our other hypothesis, that few, diverse interventions would be used, as previously found for audiology (Emanuel et al., 2011), was partly supported. For Audiology, the previous finding was confirmed, with fewer children than in other services receiving any recommended intervention. For the other services, most children received behavioral intervention recommendations that were delivered in serial therapy sessions. Interventions may have been low for Audiology because that profession is device oriented, rather than therapy focused.
Perhaps related to this, many audiologists believe that they should diagnose APD, and that SLPs should provide behavioral therapy for it (Emanuel et al., 2011). This may be appropriate, since increasing evidence suggests that LiD/APD may be specific to speech and language synthesis and understanding (Petley et al., 2021; Stewart et al., 2022).

**Relations between lab and clinical data**

We found that clinical assessments, and the intervention suggestions that followed, showed little overall relationship to the research results of the SICLiD study. The most striking difference between children with and without LiD in the SICLiD study was the consistently poorer performance of the children with LiD across the NIH Cognition Toolbox tests (Petley et al., 2021), supporting the hypothesis that LiD is primarily a cognitive impairment, rather than an auditory sensory problem (de Wit et al., 2018; Moore et al., 2010; Petley et al., 2021; Tomlin et al., 2015). In this study, however, performance on only two language tests (NIH ORR, NIH PV), among the seven NIH Toolbox subtests examined, predicted children with just a single clinical condition (ASD). These results suggest that clinicians may look beyond a child’s overall cognitive capacity during neurodevelopmental assessment.

Speech-in-noise listening has been suggested as a common problem, or even the problem experienced by children with LiD (AAA, 2010; Cameron et al., 2015; Campbell et al., 2019; Petley et al., 2021; Saunders & Haggard, 1989; Sharma et al., 2014; Vermiglio, 2014). It was therefore surprising that, in this study, we did not find a single significant relationship between three subtests of the LiSN-S and the five clinical conditions. A near-miss was the Low Cue subtest (LiSN LC) that predicted (p=0.08; d=0.47) children with a Language problem, but no research tests were predictive of APD. The general lack of relationship with the clinical conditions suggests that the LiSN-S, a test we chose primarily because of its likely relevance to everyday communication in noisy environments, may not capture well the everyday problems that children with LiD experience. Based on these findings, we predict that other commonly
used speech-in-noise tests (see (Billings, Olsen, Charney, Madsen, & Holmes, 2024) for a review) may also be unable to capture those problems.

**Caregiver reports**

Caregiver reports (ECLiPS, CCC-2) accounted for half of the predictions by SICLiD tests of clinical assessments in this study. Caregiver reports were thus better predictors of overall clinical assessments than the research tests of hearing and cognition used in SICLiD.

Previously, Bishop et al. (2006) and Bishop & McDonald (2009) showed that CCC-2 was as good a predictor of clinically assessed language impairment as widely used, standardized tests, such as the NEPSY suite for speech repetition ability (Korkman, 1998) and other tests of reading ability (Neale, 1997; Torgeson et al., 1999). Here, we also found that the CCC-2 identified clinical language problems well but, surprisingly, not as well as the SCAN, the most used clinical test for diagnosing APD in the U.S. (Emanuel et al., 2011). Also surprising was that children with a clinically identified Hearing & Processing condition, most of whom were diagnosed with APD, were predicted only by the caregiver report measures. None of the other SICLiD measures, such as speech-in-noise (LiSN-S), cognitive (NIH Toolbox), or dichotic listening ability (SCAN CW, SCAN CS, DL), frequently associated with APD in the literature (Bamiou et al., 2001; de Wit et al., 2018; Dillon & Cameron, 2021; Moore et al., 2010; Sidiras et al., 2016), predicted clinical Hearing & Processing conditions. These apparent mismatches between research and clinical findings, and the diagnostic associations with which they are assumed to be directed, were not limited to Language and Hearing & Processing, however. Both caregiver reports and the SCAN Comp also predicted clinical Attention, while standardized tests of reading and attention, together with the ECLiPS caregiver report PSS subscale, most strongly predicted ASD.

The general lack of discriminative relationship between clinical assessments and current, standard test measures may reflect misperceptions among neurodevelopmental scientists of the
true nature of clinical disorders, misperceptions among clinicians of the most important symptoms of those disorders, or both. Fortunately, adult hearing research is pivoting towards more ecological evaluations of everyday listening environments, using techniques such as ecological momentary assessment (Holube et al., 2020) and large-scale, electronic health record assessments of diverse populations (Saunders et al., 2021). Adaptations of these techniques may also provide valuable data for pediatric populations. Together, these approaches may help bridge the gap revealed here between research tests and clinical practice.

Predicting clinical assessments

Bishop & McDonald (2009) found that combining the results of a language test (NEPSY) with those of the CCC-2 led to the best discrimination between clinically referred and non-referred children, suggesting that these assessment approaches provided useful complementary information. We therefore asked whether multiple test scores might yield better prediction of the clinical conditions more generally and, indeed, that was the case. Across assessments, pairing tests represented the best compromise between minimizing collinearity, larger sample and effect sizes, and maximizing AUROC values. The end results were impressive, suggesting that the SCAN Composite, along with caregiver report measures (CCC and ECLiPS) and, for ASD, the NIH Toolbox Oral Reading Recognition test were more predictive of clinical assessments than any individual measures. Again, the SCAN, a test currently used for assessing APD, was very sensitive to Attention and Language, particularly when used together with caregiver reports. This result is logical, in that the SCAN requires quick repetition of words and sentences (language skills) in a dichotic format, in which one sentence must be attended to while the other is ignored (attention skills).

Interventions

Many therapies target functional skills in attention, fine motor, language, and sensory processing, but few behavioral therapies are designed to specifically treat listening skills. There
is a need for more research to offer evidence-based listening skill interventions. Multiple training strategies have been tried and found ineffective in generalizing to untrained skills (Halliday et al., 2012; Stewart et al., 2022; Stewart et al., 2020). Promising recent studies include de Larrea-Mancera et al. (2022), who found in college students that training broadly based auditory processing skills modestly improved a mix of supra-threshold auditory task scores relative to a control group who trained only on one task (pure tone frequency discrimination). Related cognitive skills did not differ between the groups, and long-lasting benefit was not found in the typical college student participants of this study. A second study used social speech communication training for parents of infants (Ferjan Ramírez, Lytle, & Kuhl, 2020) to show enhanced language turn-taking and skills in the infants that are preludes to enhanced communication later in life. These very different examples, employing multi-task training and very early intervention, are role models of robust design leading to secure, plausible results that may be followed up by research groups and ultimately employed by any of the services discussed in the present study. Whatever the intervention, however, the results presented here suggest that including a self- or caregiver-report measure may be a useful component of outcome evaluation.

Limitations and further work

Responses to caregiver report items in the study reported here may have reflected the (earlier) clinical assessments and may thus not have been a fair measure of association. However, the time lag between clinical visits and research testing was typically months to years, and the items in the ECLiPS and CCC-2 are, for the most part, not obviously related to specific clinical assessments.

Two SICLiD research tests, not yet discussed in detail, looked promising, but were not included in the final models. TAIL Distract, a component of a comprehensive auditory attention test for children (Zhang et al., 2012), showed a large significant influence on ASD prediction in this...
sample when paired with ECLiPS PSS and CCC SIDI (AUROC = 0.88). However, the sample size was small (n = 29) and none of the predictor p-values was < 0.08 in the model. Similarly, the DLI -15 N, a measure from the Bergen Dichotic Listening Test (Moore et al., 2020), lifted the Attention AUROC to 0.97, when used with the SCAN Comp and CCC SIDI. But, again, the sample size was small (n = 32) and the p-value exceeded 0.1. Further validation of these findings could be sought in a larger-scale study.

LiD has been classified as an umbrella symptom of other assessments (e.g. APD, DLD, ADHD, ASD) rather than as a discrete disorder (Dillon & Cameron, 2021). Although LiD is a significant issue driving caregiver concerns, as shown by the ECLiPS, it was not specifically addressed by professions other than Audiology in this study. This disconnect, between presenting concerns and the specificity of assessment and intervention, may be a reason why APD originally emerged as a diagnostic category. Unfortunately, in the absence of specific and accepted assessment and treatment approaches, children with LiD can make the rounds to different clinical services and may receive help for language, attention or psychological factors that don’t address the complex verbal communication issues encountered in daily life at home and school.

Audiologists, due to their training and experience with hearing loss and device-oriented treatment approaches, are routinely sought out where there are concerns that a child may not be hearing well. Finding normal hearing function, audiologists have sought to explain listening problems, that have no other clear explanation, through challenging tests that stress the auditory system. While these tests document the extent of listening problems in some cases, they may also reflect non-auditory system problems that are more complex, multifactorial, and not easily addressed through audioligic management. Crucially, through longitudinal assessment, we have shown that LiD has long-lasting consequences, at least into adolescence (Kojima et al., in press), and deserves to be addressed early and effectively with targeted
treatment. Seeking input to this process through wider implementation of good quality caregiver reports could alert a variety of relevant clinical professionals to the needs of children with LiD.

Conclusions

We show here that neither current research nor clinical test measures of LiD and APD are, considered separately, good predictors of clinical assessments and interventions to which they are currently directed. Caregiver report scores were more aligned with clinical practice. Some tests were predictive of clinical practice, but not in the discipline in which they are commonly used (e.g. the SCAN predicted a language diagnosis), and some test measures worked well in tandem with caregiver reports. Families of children with LiD have a wide range of concerns that extends well beyond the audiological domain. Audiologists and related professionals don’t know what to do about LiD since there is a dearth of clear evidence and agreed procedures, especially with respect to interventions. More research is needed to address these concerns and lack of knowledge, but caregiver reports provide inexpensive, simple and useful information to guide families and professionals towards appropriate outcomes.
Acknowledgements

This research was supported by Grant R01 DC014078 (awarded to DRM) from the National Institute on Deafness and Other Communication Disorders and by the Cincinnati Children’s Research Foundation. DRM is also supported by the NIHR Manchester Biomedical Research Centre. The authors gratefully acknowledge the assistance of Audrey Perdew and Nicholette Sloat in testing children, and Susan Eichert, Sandra Grether, Diala Izhiman, Katie Effler, Alexandra Parshall, and Olivia Wnek in gathering, sorting and analyzing Epic data. The authors also thank the families who supported this project through participation as well as the Summer Undergraduate Research Foundation scholars who assisted with the project.

Data Availability Statement

Data for the SICLiD study will be made available on request to the authors.
References


Table 1. Significant SICLiD predictor tests of clinical assessments. Univariate analysis.
Assessment clinical conditions (labels in bold) were Language, Hearing (and Processing), Attention, ASD and APD, although no significant predictors for APD were found. For each condition, the number of children with LiD, but without or with symptoms of that condition, is shown in the top row. Following rows show SICLiD test results (e.g. CCC GCC; see Abbreviations), scaled z-score means, number of children (n), and two sample t-test results (far right column) with probability (p) and effect size (d). (italics: 0.05 < p < 0.1, bold: p < 0.01.).

<table>
<thead>
<tr>
<th>Clinical Condition</th>
<th>SICLiD tests</th>
<th>No symptoms</th>
<th>Symptoms</th>
<th>Two sample t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Z (sd), n</td>
<td>Mean Z (sd), n</td>
<td>t(d.f.), p, d</td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCC GCC</td>
<td>-2.51 (1.39), 28</td>
<td>-3.32 (1.24), 37</td>
<td>t(63)=2.48, 0.016, 0.62</td>
<td></td>
</tr>
<tr>
<td>SCAN Comp</td>
<td>-0.11 (0.60), 27</td>
<td>-0.80 (0.85), 36</td>
<td>t(61)=3.59, 0.001, 0.92</td>
<td></td>
</tr>
<tr>
<td>SCAN FW</td>
<td>0.74 (0.64), 28</td>
<td>0.43 (0.72), 36</td>
<td>t(62)=1.74, 0.087, 0.44</td>
<td></td>
</tr>
<tr>
<td>SCAN AFG</td>
<td>0.12 (0.84), 28</td>
<td>-0.34 (0.95), 36</td>
<td>t(62)=2.02, 0.047, 0.51</td>
<td></td>
</tr>
<tr>
<td>SCAN CWD</td>
<td>-0.80 (0.83), 28</td>
<td>-1.44 (0.97), 36</td>
<td>t(62)=2.81, 0.007, 0.71</td>
<td></td>
</tr>
<tr>
<td>CCC SIDI</td>
<td>-0.15 (1.96), 28</td>
<td>1.09 (1.94), 37</td>
<td>t(63)=2.55, 0.013, 0.64</td>
<td></td>
</tr>
<tr>
<td>LiSN LC</td>
<td>-0.50 (1.11), 26</td>
<td>-1.02 (1.10), 36</td>
<td>t(60)=1.81, 0.075, 0.47</td>
<td></td>
</tr>
<tr>
<td>Hearing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECLiPS SAP</td>
<td>-2.31 (0.62), 44</td>
<td>2.56 (0.40), 21</td>
<td>t(57.04)=1.92, 0.059, 0.50*</td>
<td></td>
</tr>
<tr>
<td>ECLiPS GCC</td>
<td>-3.21 (1.24), 44</td>
<td>-2.48 (1.50), 21</td>
<td>t(63)=2.06, 0.044, 0.52</td>
<td></td>
</tr>
<tr>
<td>Attention</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCAN Comp</td>
<td>-0.13 (0.69), 21</td>
<td>0.69 (0.83), 42</td>
<td>t(61)=2.68, 0.009, 0.69</td>
<td></td>
</tr>
<tr>
<td>SCAN CS</td>
<td>-0.30 (1.02), 22</td>
<td>-0.90 (1.00), 42</td>
<td>t(62)=2.23, 0.029, 0.57</td>
<td></td>
</tr>
<tr>
<td>ECLiPS PSS</td>
<td>-1.32 (0.72), 23</td>
<td>-1.82 (0.43), 42</td>
<td>t(30.8)=3.03, 0.005, 0.84*</td>
<td></td>
</tr>
<tr>
<td>ECLiPS EAS</td>
<td>-1.48 (0.88), 23</td>
<td>-1.99 (0.73), 42</td>
<td>t(63)=2.52, 0.014, 0.63</td>
<td></td>
</tr>
<tr>
<td>CCC SIDI</td>
<td>1.44 (1.83), 23</td>
<td>0.07 (1.99), 42</td>
<td>t(63)=2.73, 0.008, 0.69</td>
<td></td>
</tr>
<tr>
<td>DL I 15N</td>
<td>-0.08 (0.68), 11</td>
<td>-0.78 (1.03), 22</td>
<td>t(31)=2.01, 0.053, 0.72</td>
<td></td>
</tr>
<tr>
<td>Autism Spectrum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECLiPS PSS</td>
<td>-1.60 (0.62), 57</td>
<td>-1.96 (0.28), 8</td>
<td>t(18.9)=2.83, 0.011, 0.75*</td>
<td></td>
</tr>
<tr>
<td>CCC SIDI</td>
<td>0.76 (1.91), 57</td>
<td>-0.88 (2.42), 8</td>
<td>t(63)=2.21, 0.031, 0.56</td>
<td></td>
</tr>
<tr>
<td>NIH PV</td>
<td>-0.28 (0.86), 52</td>
<td>0.34 (1.29), 7</td>
<td>t(57)=1.69, 0.097, 0.45</td>
<td></td>
</tr>
<tr>
<td>NIH ORR</td>
<td>-0.74 (0.77), 47</td>
<td>0.02 (1.19), 7</td>
<td>t(52)=2.25, 0.028, 0.62</td>
<td></td>
</tr>
<tr>
<td>TAIL Distract</td>
<td>-0.74 (1.60), 24</td>
<td>0.87 (0.88), 5</td>
<td>t(27)=2.16, 0.04, 0.83</td>
<td></td>
</tr>
</tbody>
</table>

*Satterthwaite (unequal variances) t-test
1"Sensory processing" diagnosis was removed from the Hearing diagnosis
2Critical p-values were B-H adjusted using the original item number (4 for language, 6 for hearing, etc.; see Table 1). Other p-values not in the test list (Table 1) were adjusted using the total item number of 16
Table 2. Best logistic models for SICLiD research predictors of clinical conditions. The first named predictor for each condition had the greater discriminatory power, indexed by the predictor probability (p). ‘n’ is the number of participants completing both predictor tasks, ‘OR’ is the odds ratio (with 95% confidence intervals) for each predictor, AUROC values show the final discriminatory power for each model (scale 0.5 – 1.0). Further details in the Table 1 caption and in the text.

<table>
<thead>
<tr>
<th>Clinical Condition</th>
<th>Predictor</th>
<th>n</th>
<th>p</th>
<th>OR, 95% CI</th>
<th>AUROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>SCAN Comp</td>
<td>63</td>
<td>0.003</td>
<td>0.28, 0.12-0.64</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>CCC SIDI</td>
<td></td>
<td>0.025</td>
<td>1.41, 1.05-1.91</td>
<td></td>
</tr>
<tr>
<td>Hearing &amp; Processing</td>
<td>CCC GCC</td>
<td>65</td>
<td>0.033</td>
<td>1.60, 1.04-2.46</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>ECLiPS SAP</td>
<td></td>
<td>0.068</td>
<td>0.36, 0.12-1.08</td>
<td></td>
</tr>
<tr>
<td>Attention</td>
<td>SCAN Comp</td>
<td>63</td>
<td>0.008</td>
<td>0.17, 0.05-0.64</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>ECLiPS PSS</td>
<td></td>
<td>0.033</td>
<td>0.40, 0.17-0.93</td>
<td></td>
</tr>
<tr>
<td>Autism Spectrum</td>
<td>NIH ORR</td>
<td>54</td>
<td>0.038</td>
<td>3.02, 1.07-8.52</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>ECLiPS PSS</td>
<td></td>
<td>0.072</td>
<td>0.11, 0.01-1.21</td>
<td></td>
</tr>
</tbody>
</table>
Figure captions

Figure 1: Auditory and cognitive SICLiD test scores of children enrolled in the study. A. Mean (± 95% CI) audiometric thresholds. B. Individual ECLiPS Total scaled score (mean typical = 10) and CCC GCC (mean typical = 100) caregiver report scaled scores with linear best fit. C. Scaled score density (violin plots in panels C-G) for the CCC GCC (left axis) and CCC SIDI (right axis; typical range -10 to 10). D. ECLiPS subscale (SAP etc) and Total scaled scores (mean typical = 10). E. SCAN-3:C subscale (FW etc, mean typical = 10) and Composite (mean typical = 100) scaled scores. F. LiSN-S subscale (LC etc) and Total Advantage scaled z-scores (mean typical = 0). G. NIH Cognition Toolbox tests (PV – OR), Fluid composite (FC), Crystalized composite (CC), Total composite (TC), Early Childhood composite (ECC). Mean typical = 100. Further details in Methods.

Figure 2: Children attending each clinical service (A – D) and assessed for each disorder (E – H). For each pie chart, the largest slice is labeled with the number of participants (n) or the percentage of all participants (%). Note that percentages did not add to 100 because many children had two or more assessments.
Figure 1

(A) Threshold (dB) vs. Frequency (kHz) for Right and Left ears.

(B) ECLIPSS vs. CCC-2 GCC.

(C) CCC2 and SIDI Score distributions.

(D) ECLIPSS Score distributions.

(E) SCAN-3:C Score distributions.

(F) LiSN-S Score distributions.

(G) NIH Toolbox Score distributions.
Figure 2

A. SERVICE ATTENDANCE

B. AVERAGE ENCOUNTERS

C. MAXIMUM ENCOUNTERS

D. INTERVENTIONS

E. CONDITIONS

F. WITHOUT SENSORY PROCESSING IN HEARING

G. APD POSITIVE

H. APD NEGATIVE