ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Predicting Persistent Developmental Stuttering Using a Cumulative Risk Approach

Cara M. Singer,[a] Sango Otieno,[b] Soo-Eun Chang,[c,d] and Robin M. Jones[e]

[a] Department of Communication Sciences and Disorders, Grand Valley State University, Allendale, MI [b] Department of Statistics, Grand Valley State University, Allendale, MI [c] Department of Psychiatry, University of Michigan, Ann Arbor [d] Department of Communicative Sciences and Disorders, Michigan State University, East Lansing [e] Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN

ARTICLE INFO

Article History:
Received March 22, 2021
Revision received July 2, 2021
Accepted September 12, 2021

Editor-in-Chief: Bharath Chandrasekaran
Editor: Julie D. Anderson

https://doi.org/10.1044/2021_JSLHR-21-00162

ABSTRACT

Purpose: The purpose of this study was to explore how well a cumulative risk approach, based on empirically supported predictive factors, predicts whether a young child who stutters is likely to develop persistent developmental stuttering. In a cumulative risk approach, the number of predictive factors indicating a child is at risk to develop persistent stuttering is evaluated, and a greater number of indicators of risk are hypothesized to confer greater risk of persistent stuttering.
Method: We combined extant data on 3- to 5-year-old children who stutter from two longitudinal studies to identify cutoff values for continuous predictive factors (e.g., speech and language skills, age at onset, time since onset, stuttering frequency) and, in combination with binary predictors (e.g., sex, family history of stuttering), used all-subsets regression and receiver operating characteristic curves to compare the predictive validity of different combinations of 10 risk factors. The optimal combination of predictive factors and the odds of a child developing persistent stuttering based on an increasing number of factors were calculated.
Results: Based on 67 children who stutter (i.e., 44 persisting and 23 recovered) with relatively strong speech-language skills, the predictive factor model that yielded the best predictive validity was based on time since onset (≥ 19 months), speech sound skills (≤ 115 standard score), expressive language skills (≤ 106 standard score), and stuttering severity (≥ 17 Stuttering Severity Instrument total score). When the presence of at least two predictive factors was used to confer elevated risk to develop persistent stuttering, the model yielded 93% sensitivity and 65% specificity. As a child presented with a greater number of these four risk factors, the odds for persistent stuttering increased.
Conclusions: Findings support the use of a cumulative risk approach and the predictive utility of assessing multiple domains when evaluating a child's risk of developing persistent stuttering. Clinical implications and future directions are discussed.

Multiple studies have explored demographic and clinical predictive factors that differentiate the approximately 80% of children who stutter who eventually recover within a few years of onset and the 20% of children who persist (for a review on the epidemiology of stuttering, see Yairi & Ambrose, 2013). However, empirical evidence is lacking on how to apply these factors to predict a child's risk for persistent stuttering (cf. Walsh et al., 2021). One approach that has been discussed within the literature, but not empirically validated, is to consider a child who presents with more predictive factors indicating persistence to be at greater risk for persistent stuttering than a child with fewer factors indicating risk (i.e., cumulative risk). To address this gap in the literature, we combined extant longitudinal data from studies conducted at Michigan State University

Correspondence to Cara M. Singer: singerca@gvsu.edu. **Special circumstances regarding authorship:** This project represents a cross-laboratory and cross-institution collaboration. Soo-Eun Chang and Robin M. Jones, senior authors, made equal contributions. ***Disclosure:*** *The authors have declared that no competing financial or non-financial interests existed at the time of publication.*

(MSU) and Vanderbilt University Medical Center (VUMC) to identify cutoff values and combinations of predictive factors that best predict a child's chances of developing persistent stuttering.

Stuttering, also known as childhood onset fluency disorder (American Psychiatric Association, 2013), is a neurodevelopmental disorder commonly characterized by disruptions in the flow of speech in the form of repetitions, prolongations, and blocks. Approximately 3%–8% of preschool-aged children meet diagnostic criteria for a stuttering disorder, with 75%–80% of these same children exhibiting natural recovery (i.e., falling below criteria for stuttering within the first several years; for a review, see Yairi & Ambrose, 2013). Predicting which children will develop persistent stuttering is one of many factors important to consider when making treatment recommendations for these young children who stutter given that persistent stuttering is associated with negative social, emotional, and vocational outcomes (Blood & Blood, 2004; Guttormsen et al., 2015; Klein & Hood, 2004). Other important considerations relative to treatment include the impact stuttering may have on the child and/or the family, such as feelings about the child's ability to communicate, the child's future, and social interactions (Guttormsen et al., 2015; Kelman & Nicholas, 2008; Langevin et al., 2010). Predicting a child's chances of developing persistent stuttering would allow for better discernment of which children should be considered for early intervention, particularly when other frank deficits (e.g., a concomitant speech or language disorder, or negative impact of stuttering) are absent. In the absence of other frank deficits, speech-language pathologists are less likely to recommend treatment for a young child who stutters (Nippold, 2004).

A flourishing area of research related to persistent stuttering has been the identification of predictive factors that can help differentiate children who eventually recover from stuttering from those who do not. Since the seminal longitudinal study conducted by Yairi and Ambrose at the University of Illinois at Urbana–Champaign in the early 1990s, multiple community-based (Kefalianos et al., 2014), multisite (e.g., Ambrose et al., 2015; Walsh et al., 2018), and single-site (e.g., Chow & Chang, 2017; Singer et al., 2019) longitudinal studies have been conducted to identify factors related to stuttering persistence. These studies made it possible for Singer, Hessling, et al. (2020) to conduct a meta-analysis to synthesize the available evidence for several predictive factors for stuttering persistence. Based on 11 longitudinal studies, this meta-analysis supported the utility of seven predictive factors for persistent stuttering: male sex, a positive family history of stuttering, older age at onset, low performance on measures of speech sound accuracy, expressive language, receptive language, and greater stuttering frequency. Examples of factors not found to be associated with stuttering persistence included frequency of specific types of disfluency, such as sound–syllable repetitions and prolongations/blocks, temperament, and receptive and expressive vocabulary, which may, at least in part, have been due to the small number of studies available for some of these analyses. Other predictive factors that have been identified within the literature, but could not be included in the meta-analysis, were time since onset (Yairi & Ambrose, 1999) and performance on a nonword repetition task (Spencer & Weber-Fox, 2014). Walsh et al. (2020) and Walsh et al. (2018) have provided evidence for the predictive value of stuttering severity and family history of persistent stuttering, respectively.

Traditional Recommendations

As the evidence linking specific predictive factors to persistent stuttering has grown, empirical studies that elucidate how to implement these predictive factors in practice have lagged. Though not empirically tested, an emphasis on the number of risk factors (i.e., predictive factors indicating risk) a child presents with can be found across published recommendations on how speech-language pathologists (Zebrowski, 1997) and parents (e.g., Guitar & Conture, 2006) might use these risk factors to evaluate a child's risk for persistent stuttering. Zebrowski (1997) developed decision "streams" in which a "child receives one point for each of the factors that he or she displays. Scores are then broadly associated with decision 'plans of actions'" (p. 24). In general, children with more risk factors are considered more likely to persist than children with fewer risk factors and are recommended a more direct therapeutic approach. Similarly, a risk factor chart developed by Ehud Yairi in Guitar and Conture (2006) suggests that caregivers consider whether their child has any of the provided six risk factors and explains that "if your child has one or more of these risk factors, you should be more concerned." Similar to Zebrowski (1997), all risk factors within the chart are given the same weight (e.g., one point). Similar approaches can be found across published tutorials and textbooks (e.g., Guitar, 2019; Logan, 2022). These types of recommendations would be considered examples of a cumulative risk approach in which all risk factors are considered individually and equally (i.e., one factor is not weighted more than a second factor) and the presence of a greater number of risk factors confers greater risk.

Empirical Support for Cumulative Risk Approaches

Cumulative risk approaches are a common method for measuring risk, especially for developmental conditions and disorders (for a review of studies that explore cumulative risk relative to childhood disorders, see Evans et al., 2013). Cumulative risk approaches "examine the number of risk [factors] experienced rather than the intensity or the pattern" of risk factors (Evans et al., 2013). The predictive factors are dichotomized using cutoff values to identify values considered "at risk" versus not at risk. In cumulative risk approaches, risk factors are considered to cumulatively influence the development of the condition or the disorder; children who present with more risk factors are considered to be at greater risk for the developmental condition or continuation of the disorder (e.g., persistence of stuttering) than children who present with fewer risk factors.

There is precedent for using cumulative risk approaches for communication disorders. For example, Hayiou-Thomas et al. (2021) found that a cumulative risk approach was a valid predictor of poor language and reading outcomes for young children. They found that 4-year-old children with three to six risk factors were more at risk to develop language or reading disorders by the age of 12 years when compared to children with one to two risk factors. Additionally, they found that including the severity of the individual risk factor did not improve predictive validity, indicating that assessing whether factors indicate risk/no risk was sufficient.

Whereas the validity of cumulative risk may seem intuitive, empirical evidence is still needed. Furthermore, cutoff values for predictive factors that indicate risk for persistent stuttering have either not been tested (e.g., 12–18 months of time since onset has been recommended by Yairi & Ambrose, 2005) or presented (e.g., stuttering frequency or speech and language scores). Identifying cutoff values is an essential step for utilizing a cumulative risk approach.

Cutoff values indicating whether risk is present or not have traditionally been identified using the lower quartile value for a given factor (Lucio et al., 2012). For example, for a predictive factor in which lower scores are considered to be associated with risk, children who performed at or below the lower quartile (i.e., at or below the 25th percentile) on the assessment would be considered to be "at risk," whereas children who perform above the lower quartile would be considered to *not* be "at risk." Similar standards (e.g., a standard score of 85 [16th percentile]) are often used in speech-language pathology to identify children with low speech and language skills (e.g., Tomblin et al., 1996; Selin et al., 2019). This method is data driven, but some have questioned whether it may "conflate rarity with severity of risk" (Evans et al., 2013, p. 42).

An alternative data-driven method for identifying cutoff values is to use a receiver operating characteristic (ROC) curve analysis in which a graph is plotted of the sensitivity and specificity for a binary outcome (e.g., identified as persistent or recovered) as the threshold of the factor is varied (e.g., different potential cutoff values). Sensitivity refers to the proportion of people with a condition (e.g., persistent stuttering) who are correctly identified; specificity refers to the proportion of people without the condition (e.g., eventual recovery) who are accurately identified. Unlike the quartile method, it is not directly associated with rarity of the scores. There is precedent for using the ROC method to identify cutoff values in the stuttering literature. Both Tumanova et al. (2014) and Walsh et al. (2020) have used the method to identify cutoff values related to stuttering frequency and stuttering severity, respectively. However, it has traditionally been used less often than the quartile method to identify cutoff values for specific factors within cumulative risk studies.

Theoretical Motivation for Exploring Cumulative Risk

Given contemporary theoretical models of stuttering, such as the dual-diathesis stressor (DD-S) model (Conture & Walden, 2012) and the multifactorial dynamic pathways (MDP) theory (Smith & Weber, 2017), exploring a cumulative risk approach is warranted. Both theories agree on two central tenets of stuttering: (a) Multiple domains are associated with the development of stuttering (e.g., biological, speech-motor, linguistic processes, and temperament), and (b) there is variability across children who stutter as to which domains, and their related skills or characteristics, influence the child's stuttering development. More specifically, in their explanation of the DD-S model, Conture and Walden (2012) explained that some children's stuttering may be attributed to language vulnerabilities, whereas other children's stuttering may be attributed to temperamental vulnerabilities or vulnerabilities in both domains. Similarly, Smith and Weber (2017) explained "a critical feature of the MDP account is an emphasis on the heterogeneity of the role of motor, language, and psychosocial factors in determining the course of this disorder [i.e., stuttering]" (p. 2497). A benefit of a cumulative risk model is that it allows for different constellations of factors instead of focusing on one explanation for why a child would be at higher risk for persistent stuttering. Furthermore, the MDP theory specifically suggests that when the child's speech-motor system is contending with multiple demands—perhaps at least partially related to the number of risk factors the child presents with—it may be more difficult for the system to produce fluent speech (i.e., for the child to naturally recover).

Recently, Walsh et al. (2021) explored an alternative approach to predicting risk for stuttering persistence motivated by similar theoretical tenets. The primary purpose of their study was to explore important relations between factors (cf. a cumulative risk model) they found to be predictive of stuttering persistence in their study sample (i.e., performance on a nonword repetition task, weighted stuttering-like disfluencies [SLDs], speech sound accuracy, and family history of stuttering). Additionally, within their study, they also explored whether a cumulative model based on all the factors was better at predicting risk for persistent stuttering than any one individual risk factor. They found that a comprehensive model yielded better predictive validity than considering any risk factor in isolation, which supports the predictive value of considering multiple predictive factors when evaluating a child's risk for persistent stuttering, a central tenet of a cumulative risk approach. Their work identifies particular relations that might assist clinicians in assessing risk for persistent stuttering, but the identification of empirically determined cutoff values and whether the presence of an increasing number of risk factors actually increases a child's chances to develop persistent stuttering awaits further exploration. The potential simplicity in which a child's risk for persistent stuttering could be evaluated and explained using a cumulative risk approach would be a strong alternative approach to the one detailed by Walsh et al. (2021).

Based on the previously described empirical support for specific predictive factors of stuttering persistence, expert clinical recommendations relative to evaluating a child's chances of developing persistent stuttering, and contemporary theoretical models of stuttering, this study aimed to evaluate whether the presence of an increasing number of predictive factors increases a child's risk for persistent stuttering.

Our primary research questions were as follows:

1. What are the optimal data-driven prognostic thresholds (i.e., cutoff values) for continuous putative predictive factors to differentiate 3- to 5-year-old children who stutter who persist from those who eventually recover based on (a) the upper or lower quartile and (b) ROC curves?
2. What is the optimal combination of predictive factors to consider when evaluating the risk for a child who stutters to develop persistent stuttering?
3. Does cumulative risk predict persistent developmental stuttering (i.e., does a child's odds of persisting increase as more predictive factors indicate persistence)?

Method

The Grand Valley State University Institutional Review Board (IRB) determined this study did not require IRB oversight due to the use of extant data; however, data transfer agreements were obtained from MSU and VUMC. Data shared originated from longitudinal prospective cohort studies previously described in the literature (e.g., MSU: Chow & Chang, 2017; Garnett et al., 2018; VUMC: Singer, Walden, & Jones, 2020; Zengin-Bolatkale et al., 2018).

The Study Sample

Eligibility

To target preschool-aged children who stutter who were followed during the period of time in which stuttering persistence/recovery was likely captured, participants from either data set were eligible for this study if they met the following criteria: (a) were between the ages of 36 and 71 months at study entry, (b) were classified as stuttering at the initial visit based on parent report and producing at least 3% SLD in either of two speech samples, and (c) were followed for at least 24 months. Due to the nature of our study, it was critical that participants had complete predictive factor data, so that the presence/absence of increased risk could be identified across all predictive factors and participants. For this reason, participants had to meet a fourth criteria of having complete predictive factor data (e.g., standardized testing related to speech sound accuracy, receptive language, expressive language, sex, family history of stuttering, age at onset, and stuttering frequency) collected during their initial visit.

Classification

Participants were classified into persisting and recovered groups based on data collected at the final visit (i.e., at least 24 months after study entry) available in both data sets—frequency of SLDs within two speech samples, stuttering severity based on the Stuttering Severity Instrument (SSI; Riley, 1994, 2009) total score, and parent report.

Parent report from the MSU data set was collected during interviews with research personnel; parent report for participants from the VUMC data set was based on scores from the Test of Childhood Stuttering Observational Rating Scale (TOCS-ORS; Gillam et al., 2009). Parent report data collected via interviews and the TOCS-ORS have previously been used as measures to determine talker group classification and found to be correlated (Tumanova et al., 2018). Participants were considered recovered if they produced less than 3% SLD across both samples, scored less than 11 on the SSI, and were reported by the caregiver to be showing near typical levels and types of disfluency (e.g., represented by a score of less than 8 on the speech fluency rating on the TOCS-ORS) at their final visit. Participants were considered exhibiting persistent stuttering if any of the three aforementioned criteria were not met.

Putative Predictive Factors

Predictive factors were selected if they were supported by empirical evidence and were available within both data sets. Seven predictive factors found to differentiate children who persist and recover based on meta-analytic evidence were included: sex, age at onset, family history of any stuttering, speech sound accuracy, receptive language, expressive language, and stuttering frequency. Time since onset was also selected as it has been found to be associated with stuttering persistence (e.g., Yairi & Ambrose, 2005) and is a commonly accepted risk factor (e.g., Clark et al., 2017; Walsh et al., 2020). Stuttering severity was included based on findings from Walsh et al. (2020) and Singer and Kelly (in press). Furthermore, although family history of persistent stuttering was not supported as a risk factor in the work of Singer, Hessling, et al. (2020), we included it due to evidence reported in the work of Walsh et al. (2018). Family history of recovered stuttering was not included due to the mounting evidence that its presence is not predictive (e.g., Singer, Hessling, et al., 2020; Walsh et al., 2021). Thus, a total of 10 predictive factors were explored.

Case and Stuttering History

Biological sex, family history, and age at onset were collected based on parent report. Participants in both studies were asked about family history of stuttering for first-, second-, and third-degree relatives (e.g., siblings, parents, aunts, uncles, cousins, grandparents, cousins, great aunt and uncles, and great grandparents). Information on whether relatives persisted was collected. At both VUMC and MSU, clinicians asked when parents first noticed their child starting to stutter and if the onset was related to any particular event or date, similar to the bracketing procedure described in Yairi and Ambrose (1992), to help caregivers narrow down the exact date that they noticed their child stuttering (i.e., age at stuttering onset). Time since onset was calculated by subtracting the parent-reported age (in months) at stuttering onset from the child's age at the initial visit.

Speech Sound and Language Skills

All speech and language skills were measured using norm-referenced standardized tests and are reported as standard scores. Speech sound accuracy was measured using the Goldman-Fristoe Test of Articulation (GFTA; Goldman & Fristoe, 2000) at both VUMC and MSU. Expressive and receptive language scores were measured using the Test of Early Language Development (Hresko et al., 1991) at VUMC. The expressive and receptive language skills for 12 of the MSU participants were measured using the Clinical Evaluation of Language Fundamentals–Preschool (Semel et al., 2004); for the remaining 22 MSU participants, their expressive and receptive language skills were measured using the Fluharty Preschool Speech and Language Screening Test–Second Edition (Fluharty, 2001). All three language tests are comprehensive language tests for preschool-age children and provide separate expressive and receptive language scores with 100 as the mean standard score and a standard deviation of 15 based on large normative samples (range: 705–2,217 children).

Stuttering Frequency and Severity

Stuttering frequency was based on video-recorded speech samples collected between the child and a clinician with expertise in stuttering and calculated based on SLDs (e.g., monosyllabic word repetitions, sound/syllable repetitions, audible prolongations, and inaudible prolongations [i.e., blocks]). At MSU, stuttering frequency counted as syllables was measured via written transcriptions using Computerized Language Analysis (https://dali.talkbank.org/clan/). At VUMC, stuttering frequency counted as words was measured via real-time coding and converted to syllables using a conversion factor (Yaruss, 2000) that is suitable for children in this age range. Real-time measurement of stuttering frequency is a valid measure of stuttering frequency (O'Brian et al., 2013) with very high reliability (Tumanova et al., 2014) that shows robust correlation with stuttering frequency from written transcriptions (Yaruss, 1998). All stuttering frequency measures are reported as percent syllables stuttered. Stuttering severity was based on the SSI total score.

Analytical Plan

Our analytical plan can be described in three main steps: (1) identifying cutoff values for our evidence-based predictive factors, (2) exploring and identifying the optimal combination of predictive factors that best differentiates children who persist and recover, and (3) identifying whether an increasing number of predictive factors indicating persistence is associated with greater risk of later stuttering persistence.

Identifying Cutoff Values

Because both the quartile and ROC method have been used to identify cutoff values within the literature and cutoff values play a fundamental role within a cumulative risk approach, we used both methods for comparison purposes.

To identify prognostic cutoff values based on the quartile method (RQ1a), the cutoff values for each continuous predictive factor was identified by calculating the quartiles based on all participant data. For putative predictive factors in which higher values are considered to be associated with greater chances of persistence compared to lower values (i.e., age of stuttering onset, time since stuttering onset, stuttering severity, and stuttering frequency), the upper quartile (i.e., score at the 75th percentile) was considered to be the cutoff value. For putative risk factors in which lower scores are considered to be associated with greater chances of persistence compared to higher scores (i.e., speech sound accuracy, receptive language, and expressive language), the lower quartile (i.e., score at the 25th percentile) was considered to be the cutoff value.

To identify prognostic cutoff values based on the ROC method (RQ1b), the optimal cutoff value for each continuous predictive factor was identified based on Youden's method (Youden, 1950) using the *OptimalCutpoints* package (Lopez-Raton & Rodriguez-Alvarez, 2019) in R (R Core Team, 2013). This package conducts ROC analyses and compares the sensitivity and specificity of possible prognostic thresholds. Persistence was modeled as the outcome of interest. The directionality of the model differed across putative predictive factors based on whether higher or lower scores have been found to be associated with greater risk for persistent stuttering.

Applying Cutoff Values

Once the cutoff values were identified, they were applied to our data set to determine whether a participant's score was indicative of persistence. For each factor, participants whose score was considered to indicate persistence (i.e., meet or exceed the cutoff value) received a score of "1"; participants whose score fell below the cutoff value received a "0." Each of the seven continuous factors (i.e., speech sound accuracy, expressive language, receptive language, stuttering frequency, stuttering severity, time since stuttering onset, and age at onset) was assessed twice due to the use of two cutoff values (one using the quartile method, one using the ROC method). For the three binary predictive factors (i.e., sex, family history of any stuttering, and family history of persistent stuttering), the decision of which category indicated persistence was based on empirical findings (e.g., Singer, Hessling, et al., 2020). Male sex was given a score of "1"; participants who had a known family history were given a score of "1." Odds ratios (*OR*s) were then conducted to identify how well each binary predictive factor was related to eventual persistence.

Comparing Predictive Models

To compare all possible models based on the putative predictive factors (RQ2), we used all-subsets regression. Using the *leaps* (Lumley, 2020) package in R, we compared all possible models based on an increasing number of factors (i.e., models with one factor up to models with 10 risk factors). For each method, models based on the same number of risk factors (e.g., combinations of seven risk factors) were compared using Mallow's C_p statistic (Mallows, 1973). The model in which C_p is closest to p + 1 where p is equal to the number of included risk factors was considered to be best fit. Because all possible combinations of the factors are evaluated, multicollinearity between independent factors is not as much a concern as other types of regression (Kraha et al., 2012). The combined effect of the factors is more relevant than any one factor's individual contribution. All-subsets regression was conducted separately for combinations of factors based on cutoff values from the quartile and ROC methods.

The best-fit models for the resulting 20 models (i.e., 10 models per method) were then compared using ROC curves. The area under the curve (AUC), which is computed with ROC curve analyses, is a measure of a model's predictive validity: Values below .7 are considered "nonuseful," .7–.79 are considered "fair," .80–.89 are considered "good," and values above .9 are considered "excellent" (Carter et al., 2016). Measures of diagnostic validity (i.e., sensitivity, specificity, positive predictive value, and negative predictive value) were calculated to further explore predictive validity. Figure 1 depicts the 2 × 2 contingency table of predicted and true outcomes and related formulas used to derive these diagnostic values.

Evaluating a Cumulative Risk Approach

Binomial logistic regression was used to identify the amount of risk associated with each additional predictive factor indicating persistence (RQ3). Persistence was entered as the dependent variable; the number of predictive factors was the explanatory variable.

Results

Descriptive Statistics

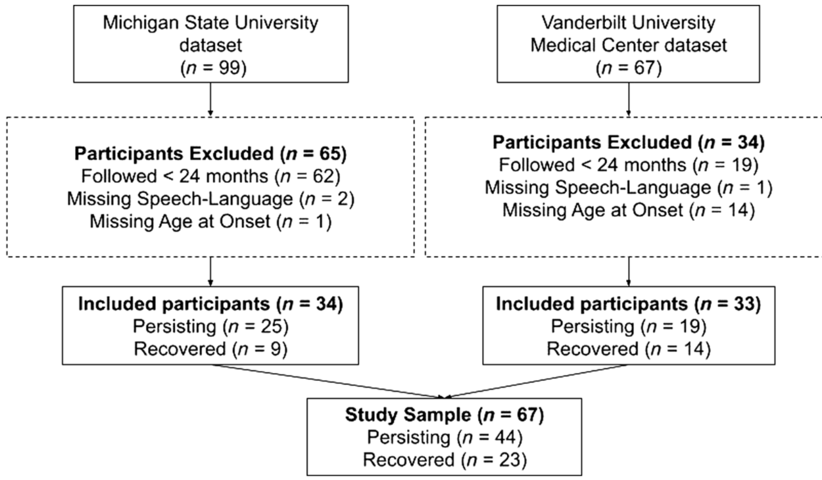
Based on the predetermined eligibility criteria, 67 participants (i.e., 44 persisting; 23 recovered) were included in this study sample (see Figure 2). Participants were followed for an average of 31 months (range: 24–55 months). Participants at VUMC were seen every 8–10 months (i.e., up to five visits per child); participants at MSU were seen yearly (i.e., up to four visits per child). By the final visit, 50% of the children who exhibited persistent stuttering were at least 4 years post stuttering onset; 82% were at least 3 years post onset. The MSU participants were slightly older than the VUMC participants, $t(65) = 3.17$,

Figure 1. A 2 × 2 contingency table and formulas used to derive diagnostic values.

	$\text{Persisting}_{\text{Outcome}}$	$\text{Recovered}_{\text{Outcome}}$
$\text{Persisting}_{\text{Predicted}}$	a	b
$\text{Recovered}_{\text{Predicted}}$	c	d

Sensitivity = a/(a + c)
Specificity = d/(b + d)
Positive Predictive Value = a/(a + b)
Negative Predictive Value = d/(c + d)

Figure 2. Flow chart depicting study sample.



p = .002; no other significant differences were found between the two samples related to sex, maternal education, stuttering frequency, age at stuttering onset, time since stuttering onset as assessed at the initial evaluation, receptive vocabulary, or expressive vocabulary. Descriptive statistics for the two samples can be found in Appendix A.

Demographics of the persistent and recovered groups are reported in Table 1. Children with persistent stuttering, as a group, entered the study at an older age and had been stuttering for longer than children who recovered, despite similar ages of stuttering onset. No other significant differences were found between the two groups related to sex, family history of any stuttering, family history of persistent stuttering, maternal education, age of onset, stuttering frequency or severity, speech sound accuracy, receptive or expressive language, or study length.

Cutoff Values for Continuous Risk Factors

Identifying Cutoff Values (RQ1)

Calculated cutoff values for continuous predictive factors based on both the quartile and ROC methods can be found in Table 2. Comparisons of our lower quartile values for the speech and language scores to the expected lower quartile values based on a normal distribution (i.e., a standard score of 90) indicate that our study sample's skills are higher than the general population's. *OR*s were calculated to determine whether each binary predictor was associated with later persistence and are reported in Table 2. Statistical significance is indicated by the 95% confidence interval (CI) not crossing a value of 1. None of the *OR*s for the quartile-based cutoff values reached statistical significance. *OR*s for speech sound accuracy, receptive language, and time since onset based on ROC-based cutoff values were significant. No

Table 1. Study sample demographics organized by persistent and recovered groups.

Demographic variable	Recovered (*n* = 23)	Persistent (*n* = 44)	Comparison	*p*
Sex	16 M; 7 F	32 M; 12 F	$\chi^2(67) = 0.07$	.785
Family Hx of stuttering	10 (–); 13 (+)	23 (–); 21 (+)	$\chi^2(67) = 0.47$	.670
Family Hx of P. stuttering	19 (–); 4 (+)	34 (–); 10 (+)	$\chi^2(67) = 0.26$	.610
Age (months)	44.3 (6.3)	49.1 (8.3)	$t(65) = 2.64$	**.011**
Maternal education	6.2 (0.74)	6.3 (0.8)	$t(65) = -0.02$	.987
Age of onset (months)	32.9 (7.7)	32.9 (9.8)	$t(65) = -0.02$	.986
Time since onset (months)	11.4 (6.6)	16.2 (9.8)	$t(65) = -2.09$	**.040**
Percent syllables stuttered	6.6 (3.5)	6.6 (3.8)	$t(65) = 0.07$	.942
SSI total score	18.6 (6.6)	18.6 (5.1)	$t(65) = 0.02$	.987
GFTA (standard score)	110.4 (16.1)	106.0 (9.1)	$t(65) = 1.23$	.228
Receptive language (standard score)	114.4 (17.2)	109.3 (16.9)	$t(65) = 1.15$	.257
Expressive language (standard score)	113.61 (16.3)	106.3 (13.8)	$t(65) = 1.85$	.072
Study length (months)	29.9 (5.8)	31.7 (8.2)	$t(65) = -1.01$	.319

Note. All data except for study length were collected at study entry. Maternal education based on Hollingshead (1975). Bolded values are significant at $p < .05$. M = male; F = female; Hx = history; P. stuttering = persistent stuttering; (–) indicates the absence of family history; (+) indicates the presence of family history; SSI = Stuttering Severity Instrument; GFTA = Goldman-Fristoe Test of Articulation.

Table 2. Cutoff values and related odds ratios (*OR*s) and 95% confidence intervals (CIs) based on the quartile and ROC methods.

Putative risk factor	Quartile			ROC		
	Cutoff value	*OR* [95% CI]	*p*	Cutoff value	*OR* [95% CI]	*p*
Speech sound (standard score)	103	1.86 [0.58, 6.00]	.312	**115**	**13.0 [3.48, 48.55]**	**< .001**
Receptive language (standard score)	98	0.86 [0.28, 2.60]	.785	**127**	**5.01 [1.43, 17.54]**	**.012**
Expressive language (standard score)	97	1.32 [0.43, 4.08]	.647	106	2.83 [0.94, 8.53]	.066
Age of onset (months)	39	0.94 [0.30, 3.00]	.915	45	2.70 [0.53, 13.71]	.307
Time since onset (months)	21	3.11 [0.79, 12.24]	.101	**19**	**3.61 [1.05, 12.38]**	**.037**
Stuttering severity (SSI total score)	22	0.80 [0.28, 2.29]	.689	17	1.77 [0.63, 4.96]	.289
Stuttering frequency (% syllables stuttered)	8.7	0.67 [0.22, 2.09]	.505	2.3	4.10 [0.35, 47.77]	.306

Note. Bolded cutoff values yield a statistically significant *OR*. ROC = receiver operating characteristic; SSI = Stuttering Severity Instrument.

statistically significant association was found between persistence and sex (*OR* = 1.17, 95% CI [0.39, 3.54], *p* = .783), family history of any stuttering (*OR* = 0.70, 95% CI [0.26, 1.94], *p* = .509), or family history of persistent stuttering (*OR* = 1.40, 95% CI [0.39, 5.07], *p* = .637).

Predictive Models

Exploring Factor Combinations (RQ2)

Findings from the all-subsets regression analyses are reported in Table 3. The combination with k number of factors, from one to all 10 factors, that yields the best discrimination is identified when either the quartile- or ROC-based cutoff values are applied. For example, when considering all possible combinations of three factors, the best 3-factor model based on the ROC method included speech sound accuracy, time since onset, and stuttering severity. To compare the diagnostic validity between the models, AUCs are provided.

The model with the highest AUC (AUC = .819; considered “good”) was based on the ROC method and included six factors (i.e., 6-factor model): speech sound accuracy, time since onset, stuttering severity, expressive language, receptive language, and age at onset. Notably, the models with four and five factors (i.e., 4- and 5-factor models, respectively) based on the ROC method also yielded high values (AUC = .815). Figure 3 depicts the number of predictive factors indicating persistence that each of our 67 participants exhibited when considering the (a) 4-factor model, (b) 5-factor model, and (c) 6-factor model using the ROC-based cutoff values. As depicted in the figures, the presence of two factors indicating persistence in the 4-factor model and the presence of three factors indicating persistence in the 5- and 6-factor models best differentiate children who eventually persisted and recovered.

Table 4 reports the diagnostic validity values and results from the binary regressions of the 4-, 5-, 6-factor models based on the ROC cutoff values (see Table 3) to compare the utility of these models as potential screeners for persistent stuttering. For each model, we identify the cutoff value, or the number of factors indicating persistence

Table 3. Results from the all-subsets logistic regression based on both cutoff methods.

k	ROC models		Quartile models	
	Predictive factors	AUC	Predictive factors	AUC
1	SS	0.737	TSO	0.594
2	SS, RecL	0.773	SS, TSO	0.626
3	SS, TSO, StS	0.790	SS, TSO, FH_any	0.584
4	SS, TSO, StS, ExpL	0.815	SS, TSO, FH_any, SF	0.555
5	SS, TSO, StS, ExpL, RecL	0.815	SS, TSO, FH_any, SF, FH_P	0.554
6	SS, TSO, StS, ExpL, RecL, Onset	0.819	SS, TSO, FH_any, SF, FH_P, ExpL	0.565
7	SS, TSO, StS, ExpL, RecL, Onset, FH_P	0.791	SS, TSO, FH_any, SF, FH_P, ExpL, RecL	0.553
8	SS, TSO, StS, ExpL, RecL, Onset, FH_P, Sex	0.783	SS, TSO, FH_any, SF, FH_P, ExpL, RecL, Onset	0.563
9	SS, TSO, StS, ExpL, RecL, Onset, FH_P, Sex, FH_any	0.746	SS, TSO, FH_any, SF, FH_P, ExpL, RecL, Onset, Sex	0.551
10	SS, TSO, StS, ExpL, RecL, Onset, FH_P, Sex, FH_any, SF	0.751	SS, TSO, FH_any, SF, FH_P, ExpL, RecL, Onset, Sex, StS	0.536

Note. The best predictive models based on an increasing number of factors (k) are reported. ROC = receiver operating characteristic; AUC = area under the curve; SS = speech sound accuracy; RecL = receptive language; ExpL = expressive language; TSO = time since onset; StS = stuttering severity; Onset = age at onset; FH_P = family history of persistent stuttering; FH_any = family history of any stuttering; SF = stuttered frequency.

Figure 3. Number of predictive factors indicating persistence that study participants presented with based on (a) the 4-factor model that includes speech sound skills, time since onset, stuttering severity, and expressive language; (b) the 5-factor model that includes speech sound skills, time since onset, stuttering severity, expressive language, and receptive language; and (c) the 6-factor model that includes speech sound skills, time since onset, stuttering severity, expressive language, receptive language, and age at onset. ROC = receiver operating characteristic.

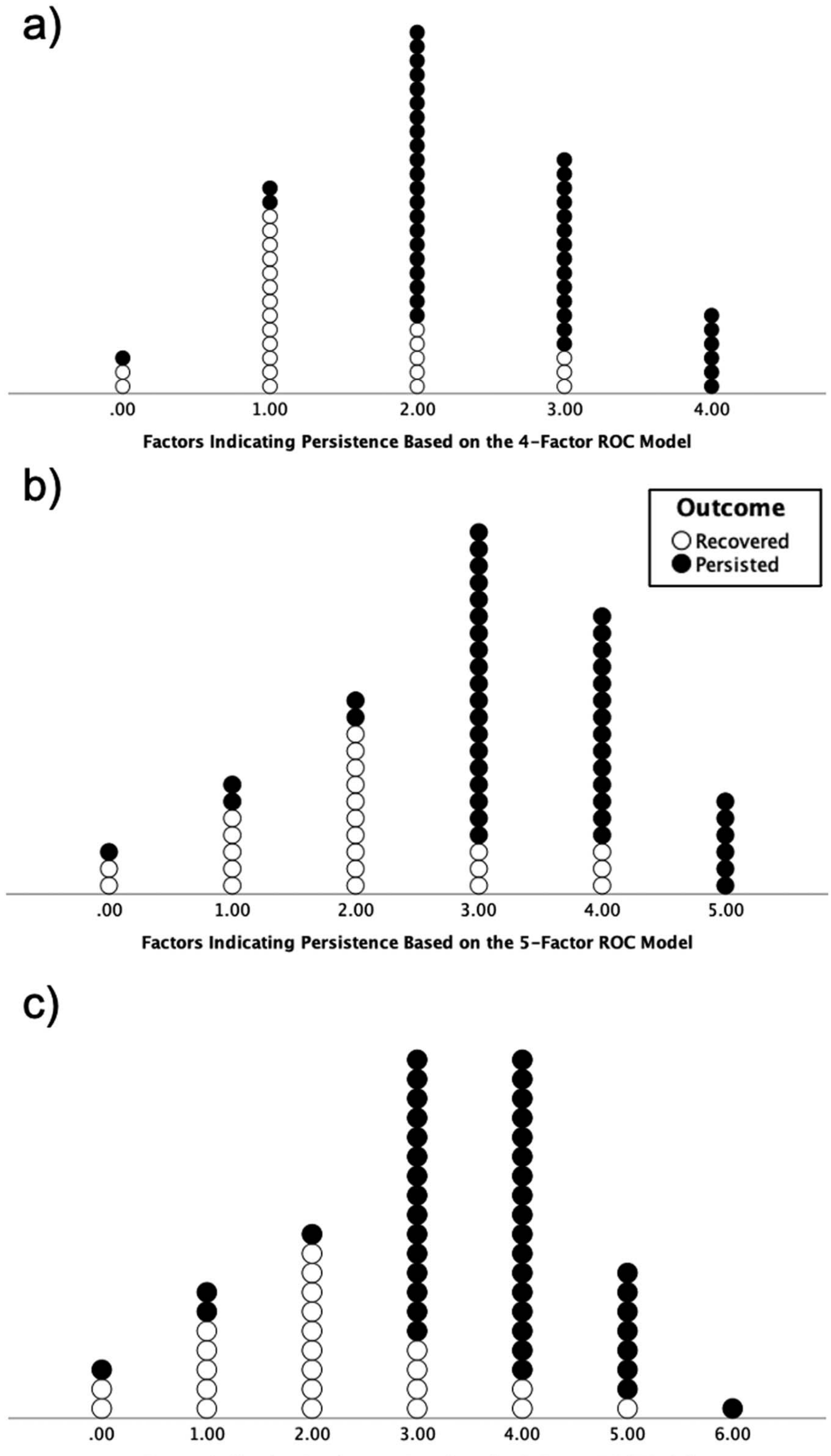


Table 4. Diagnostic validity and odds ratios (*OR*s) with 95% confidence intervals (CIs) for select models.

Model	Cutoff	Sensitivity	Specificity	Positive predictive value	Negative predictive value	*OR*s	
						Estimate	95% CI
4-factor	2 factors	93	65	84	83	4.7	[2.1, 10.7]
5-factor	3 factors	89	74	83	77	3.0	[1.7, 5.3]
6-factor	3 factors	91	73	85	80	2.8	[1.6, 4.7]

that confers elevated risk to develop persistent stuttering, and the associated diagnostic validity based on the cutoff value. For example, when using the 4-factor model as a screening tool and applying the cutoff value of at least two factors indicating persistence, the screener accurately identified 93% of children who persisted (i.e., sensitivity) and 65% of children who recovered (i.e., specificity). The 4-factor model yielded the highest sensitivity, but lowest specificity. The lower specificity is due to accurately identifying one less child who recovered than the other models. The 4-factor model accurately identified one additional child who persisted than the 6-factor model. Given our sample size (i.e., 44 persistent children; 23 recovered children) the diagnostic validity values are vulnerable to these small differences.

Evaluating Cumulative Risk (RQ3)

The *OR*s were calculated using binomial logistic regression. To evaluate the multicollinearity for the three regression models, we calculated the variance inflation factor (VIF) for each predictor within each model using the *performance* (Lüdecke et al., 2021) package in R. All VIFs were < 1.5, indicating low correlation of each predictor with the other predictors (Thompson et al., 2017). The 4-factor model yielded the highest *OR*, which indicated that for each additional predictive factor indicating persistence, a child's odds to persist increases by 4.7 (p = .012). The *OR*s across the four models were all significant, indicating the odds to persist increase with the presence of additional predictive factors indicating persistence.

Exploratory Analyses

Associations Between Predictors

To explore relations between our binary predictors based on the ROC method, Fisher exact tests were conducted. Significant values indicate that both predictive factors are likely to indicate the same outcome (i.e., persistence or recovery). Children with a family history of persistent stuttering were likely to exhibit receptive language skills and a time since onset value that also indicated persistence. Children whose receptive language skills indicated persistence were likely to also exhibit speech sound accuracy and expressive language skills indicative of persistence. Full results are reported in Table 5.

Participant Profiles

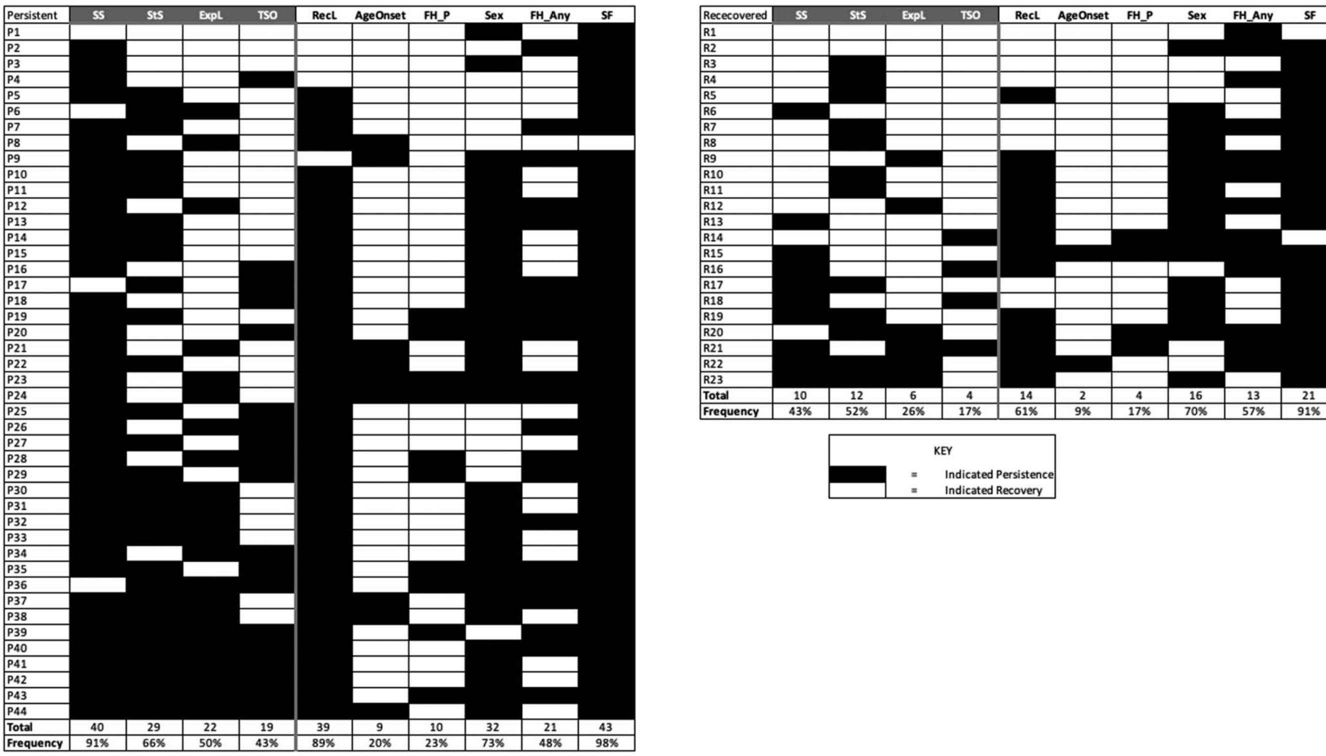
We report the predictive factors indicating persistence (black) versus recovery (white) exhibited by our 67 participants in Figure 4; the factors included in the 4-factor model are outlined in gray. Participants are reported in order of increasing number of predictive factors indicating persistence. The number and frequency of children within each group who were found to be "at risk" for a given factor are reported in the last two rows. Upon visual analysis of Figure 4, we identified that a large proportion of the girls who persisted were found toward the top of the table (i.e., they presented with fewer factors indicating risk) and that children with a family history of persistent stuttering were found toward the bottom of the tables (i.e., they presented with more factors indicating

Table 5. The *p* values for the relations between ROC-based binary predictive factors (persistence/recovery) using Fisher's exact tests.

Variable	TSO	SS	RecL	ExpL	SF	StS	FH_any	FH_P	Sex
Age onset	.082	.053	.435	**.042**	.421	.738	1.00	.686	.715
TSO		.281	.999	.999	0.324	1.20	.305	**.012**	.169
SS			**.034**	.268	.156	.781	.262	.999	.999
RecL				**< .001**	.511	.133	.560	**.031**	.517
ExpL					.999	.999	.460	.231	.784
SF						.054	.999	.511	.192
StS							.212	.368	.785
FH_any								**< .001**	.590
FH_P									.999

Note. Bolded values are significant at $p < .05$. TSO = time since onset; SS = speech sound accuracy; RecL = receptive language skill; ExpL = expressive language; SF = stuttered frequency; StS = stuttering severity; FH_any = family history of any stuttering; FH_P = family history of persistent stuttering.

Figure 4. Predictive factors indicating persistence across all study participants based on the ROC cutoff values. Black indicates persistence; white indicates recovery. SS = speech sound accuracy; StS = stuttering severity; ExpL = expressive language; TSO = time since onset; RecL = receptive language; FH_P = family history of persistent stuttering; FH_Any = family history of any stuttering; SF = stuttering frequency. ROC = receiver operating characteristic.



risk). Based on these observations, we conducted two post hoc analyses. Fisher's exact tests with Monte Carlo estimation supported our observations that (a) children with family history of persistent stuttering exhibit more factors indicating persistence than children without (p = .003) and (b) females are likely to persist with fewer predictive factors indicating persistence than males (p = .02).

Discussion

This study provides initial evidence to support a cumulative risk approach to predict stuttering persistence. We explored 10 putative predictive factors for stuttering persistence based on past empirical evidence (e.g., Singer, Hessling, et al., 2020; Walsh et al., 2020; Yairi & Ambrose, 2005) and identified cutoff values for continuous factors to determine values indicating elevated risk for persistent stuttering. We then identified a combination of predictive factors that yielded the best predictive values for our study sample as well as cumulative risk (i.e., increased odds of persistence with increased number of predictive factors). These results and their clinical implications are discussed below.

Optimal Factor Combination and Cumulative Risk

Optimal Combination of Predictive Factors

We consider the 4-factor model based on ROC cutoff values, which includes speech sound accuracy, expressive language, time since onset, and stuttering severity, to represent the most clinically useful screener for stuttering persistence based on our sample. Although the 6-factor model yielded a slightly higher AUC, one measure of predictive validity, the 4-factor model yielded a greater sensitivity (93%) than the 6-factor model and adequate specificity (65%). When considering screeners for persistent developmental stuttering, Walsh et al. (2020) identified that sensitivity is often prioritized over specificity because "failing to identify a true positive [i.e., a child who will persist] could have profound ramifications. . .[whereas] recommending treatment for a child who would have recovered without it. . .may be a more acceptable trade-off with less impactful consequences." (pp. 2562–2563). Therefore, the 4-factor model represents a better screener than the 6-factor model despite the latter's higher AUC (see Tables 3 and 4 for comparisons).

These factors within the 4-factor model represent diverse aspects of a child: (a) stuttering history (i.e., time since onset), (b) stuttering behaviors (i.e., stuttering severity), (c) speech sound production skills (i.e., speech sound accuracy), and (d) language skills (i.e., expressive language skills). These domains have commonly been attributed to the development of persistent stuttering (e.g., Smith & Weber 2017) and are similar to the factors included in Walsh et al.'s (2021) comprehensive model: family history of (any) stuttering, weighted SLD, speech sound accuracy, and performance on a nonword repetition task. Despite the different samples and measures, both studies identified overlapping domains (e.g., stuttering behaviors, speech sound production skills, aspects of language) from a larger set of predictors. Whereas there is large variability between which risk factors are related to persistent developmental stuttering across study samples (Singer, Hessling, et al., 2020), our findings and the findings from the work of Walsh et al. (2021) provide converging lines of evidence on the domains that are helpful to assess in predicting a child's stuttering persistence.

Cumulative Risk

In support of the cumulative risk approach, we found that for each additional predictive factor that indicated persistence in the 4-factor model, a child's odds to persist increased nearly fivefold. This result provides compelling evidence that risk of persistence dramatically increases with the presence of additional risk factors. We also found that the presence of at least two factors indicating persistence confers elevated risk to develop persistent stuttering. These findings support the tenet of the MDP theory that a combination of factors, not one factor in isolation, influences whether a child develops persistent developmental stuttering. As seen in Figure 4, the children who persisted were heterogeneous in that they exhibited various combinations of these four predictive factors. This heterogeneity highlights the importance of a screener to be flexible and comprehensive.

Exploring Methods for Identifying Cutoff Values

To identify whether a predictive factor indicates persistence, we calculated cutoff values. Previous tools for evaluating risk for persistent stuttering (e.g., the Stuttering Foundations Risk Factor Chart) reported that "low" speech and language scores are indicative of risk, but no clear cutoff value had previously been identified. We explored two methods of identifying cutoff values and found those based on the ROC method yielded better predictive validity than those based on the quartile method, the latter being the more standard practice. All models based on the quartile method were considered "nonuseful," whereas models based on the ROC method were considered "fair" or "good" based on the criteria described by Carter et al. (2016). Similar to our results, Plante and Vance (1994) also found that their empirically based cutoff scores for standardized language tests yielded better discrimination between children with and without specific language impairment than "standard" cutoff values (e.g., 1–2 *SD*s below the norm).

The overlapping and nonlinear performance between children who persisted and recovered may have resulted in the quartile method being less effective than the ROC method. This overlap in performance can be observed in Appendices B–H, which depict the individual values of our participants across the seven continuous predictive factors, as evidenced by similar boxplots between both groups. Given this performance, it is understandable that more "typical" or "frequent" scores, like those identified using the quartile approach, may not sufficiently differentiate the groups. Similarly, Plante and Vance (1994) explained that their identified cutoff scores were higher than the standard values for two of their explored assessments because of the overlap in performance between children with and without specific language impairment. Additionally, the boxplots also indicate a nonnormal distribution across some factors as indicated by the bolded median line being closer to the top or bottom of the box and whiskers of unequal lengths (e.g., receptive language for the recovered group and percent syllables stuttered for the persistent group). Normality may influence results using a central tendency approach, like the quartile method. Hayiou-Thomas et al. (2021), for example, altered their method of identifying cutoff values based on the skewness of individual factors. They used a criterion of -1.25 *SD* for data with normal distributions and the 10th percentile for the factor (i.e., letter knowledge) with high skewness. The ROC method can be consistently applied to data regardless of the distribution of the data.

Individual Cutoff Values

For some factors, our identified cutoff values based on the ROC method aligned with previous research and speculation. For example, the cutoff values for time since onset (i.e., $\geq$ 19 months) and age at onset (i.e., $\geq$ 45 months) align with previous estimates from the literature (e.g., Yairi & Ambrose, 2005). However, the cutoff values for speech and language skills and stuttering frequency could be considered more unexpected.

The stuttering frequency cutoff value (i.e., $\geq$ 2.3% syllables stuttered) is below the typical gold standard of 3% SLD; the speech sound (115 standard score) and receptive language (127 standard score) cutoff values are both considered "above average." The potential for participants to score lower than 3% SLD within a speech sample at their initial visit and be considered stuttering (due to consideration of other factors, such as SSI scores and expressed parent concern of stuttering) is a probable contributing factor to the lower-than-expected %SLD cutoff value. Additionally, our sample's high maternal education (i.e., most mothers had completed college or graduate degrees) and high speech and language scores may have influenced the related higher-than-expected cutoff values. However, this potential bias is common among many longitudinal studies of stuttering whose participants have had high speech and language skills due, in part, to recruiting participants from research university communities and surrounding areas (e.g., Ambrose et al., 2015; Yairi, Ambrose, Paden, & Throneburg, 1996). In fact, Singer, Hessling, et al. (2020) found that the mean receptive language standard score of 167 children who recover, reported across multiple studies, was 123 (standard deviation of 17).

Lastly, there has been no previous evidence to suggest that "clinically significant" low scores are needed for a child to be "at risk" for persistent stuttering. Previous findings have indicated that children with low speech and language abilities may be more likely to persist and children with high speech and language abilities may be more likely to recover (e.g., Ambrose et al., 2015; Spencer & Weber-Fox, 2014; Yairi, Ambrose, Paden, & Throneburg, 1996). Our findings extend the relation between speech and language skills and persistence by indicating that children with average speech and language skills may be more appropriately grouped with the children at greater odds to persist than to recover. In other words, average speech and language skills are not necessarily indicative of eventual recovery. Finally, it is worth emphasizing that the presence of an average performance on a test of speech sound accuracy or language, when it is the isolated factor indicating persistence, is not predictive of persistence on its own; at least two factors need to indicate persistence for a child to be at elevated odds for persistent stuttering.

Speech and Language Skills

The cutoff value of a 115 standard score on the GFTA was not only included in the 4-factor model, but it was also the factor that discriminated between children who persist and recover the best on its own. Children with standard scores 115 and below were estimated to be 13 times (95% CI [3.48, 48.55]) more likely to develop persistent stuttering than children with above average scores (> 115). Relatedly, Walsh et al. (2021) found that higher speech sound accuracy increased a child's chances of recovering. The convergence in these findings is important and speaks to the generalizability of the results given that the same conclusion (i.e., higher speech sound accuracy predicts recovery) were reached despite differences in sample characteristics of the two studies: The study sample from Walsh et al. exhibited speech sound accuracy scores that were lower than in the general population, whereas the opposite was observed in the current sample.

Within the speech and language domain, a score at or below 106 standard score on an expressive language measure (considered "average") was also found to be indicative of persistence. Although language scores of children who eventually persist and recover have been found to decrease across time (i.e., regress toward the mean), the scores of children who recover have commonly remained higher (Ambrose et al., 2015; Yairi, Ambrose, & Cox, 1996). Interestingly, receptive language, with a cutoff value of 127, was a more predictive indicator of persistence than expressive language when compared individually, but only expressive language was kept in the optimal, 4-factor model. This may be explained by our finding that children whose receptive language skills indicate persistence are likely to also exhibit weaker speech sound accuracy that indicates persistence; no such relation was found between speech sound accuracy and expressive language. Expressive language may identify a unique subset of children whose elevated risk for persistent stuttering may have gone undetected. In support of this speculation, when receptive language was added (i.e., the 5-factor model), sensitivity decreased.

Children whose expressive language skills indicate persistence were also found to exhibit an age of onset that indicates persistence (i.e., they began stuttering after 45 months of age). The relation between these factors likely limited the unique predictive validity age at onset contributes; age at onset was also not included in the optimal model.

Time Since Onset

As previously mentioned, our cutoff value for time since onset (i.e., $\geq$ 19 months) aligns with previous literature. Yairi and Ambrose (1992) identified a differentiation between the trajectories of children who eventually persist and recover occurs by the time children are 18–20 months post onset. As has been described by Smith and Weber (2017), perhaps children's speech-motor systems learn how to contend with their stressors (e.g., low speech or language skills) and eventually produce fluent speech (i.e., recover) when there is enough time during the critical developmental period. Perhaps our ~1.5-year cut-point is reflective of that time window.

Stuttering Severity

The cutoff value for stuttering severity (i.e., $\geq$ 17 total score on the SSI) is reflective of a "moderate" stuttering severity. Findings indicated that children who present with moderate or severe stuttering behaviors, based on the SSI, are more likely to persist than children with mild stuttering behaviors, particularly in the presence of another factor indicating persistence. In fact, all seven of the participants whose only indication of persistence was stuttering severity (out of the four predictors) went on to eventually recover. Similarly, Walsh et al. (2020) identified that a moderate severity, based on weighted SLD, increased a child's odds to persist especially in the presence of other risk factors (i.e., sex and family history).

Yairi and Ambrose (1999) found that while the stuttering behaviors of both children who persist and recover decline over time, the largest decline in stuttering frequency of children who eventually recovered was observed within the first 18 months of stuttering, which is representative of our time since onset cutoff value. Stuttering frequency, which is a dimension reflected within the stuttering severity rating, was not included in the optimal model. This is unsurprising considering that only four participants presented with frequencies indicative of recovery given the low cut-point (i.e., 2.3% syllables stuttered). Walsh et al. (2021) also found that stuttering frequency was not predictive within their set of factors.

Family History and Sex

Although sex and family history are two of the most well-documented (Singer, Hessling, et al., 2020) and clinically valued (Singer & Kelly, in press) predictive factors, they did not add predictive value when the other four factors were considered. Given that 32% of persistent children were female and 52% of our sample had known family histories of stuttering, which align with previously reported samples (e.g., sex: Ambrose et al., 2015; family history: Yairi, Ambrose, Paden, & Throneburg, 1996), we cannot attribute these findings to sampling bias. Instead, our post hoc analyses revealed intriguing relationships involving family history and sex that may help explain why both factors were not included in the optimal model.

Family History

We found that children with family history of persistent stuttering exhibit more predictive factors than children without. Considering the genetic underpinnings of stuttering (e.g., Kraft & Yairi, 2012) and speech and language skills (Kang & Drayna, 2011; Peterson et al., 2007), it may be that when the child inherits stuttering, the child also inherits other factors that contributed to the family member's persistent stuttering (e.g., delayed or dissociated speech and language skills; Conture & Walden, 2012). Studies have shown a link between the presence of a family history of stuttering and both attention-deficit/hyperactivity disorder and lower language scores (Choi et al., 2018; Donaher & Richels, 2012). Relatedly, we found that children with family history of persistent stuttering were likely to also have receptive language skills and a time since onset that indicated persistence.

Whereas our findings do not rule out family history as a predictor of persistence, they provide a new perspective on its predictive value relative to other factors with greater predictive utility. When considering the models based on the quartile method cutoff values (see Table 3) family history was included in more models. This finding supports its predictive value especially when other less predictive factors and their related cutoff values are being used.

Sex

A post hoc analysis revealed that females who persist exhibit fewer predictive factors indicating persistence than males who persist. Whereas this finding does not suggest that females are at greater risk to persist than males, it suggests that perhaps this study did not include predictive factors that most influence females who persist or, alternatively, females who stutter are more vulnerable to the effects of the explored predictive factors than males. Another possibility is that there may be unexplored unique biological factors that confer risk for persistent stuttering in girls, including genetics and neural anatomy. Anecdotal evidence points to the possibility of greater genetic loading for females who stutter (Ambrose et al., 1993; Kidd, 1984; Kidd et al., 1981). Females, who have been found to have generally greater interhemispheric connectivity than males (Ingalhalikar et al., 2014), may be more likely to be able to compensate for the left hemispheric deficit commonly found in people who stutter (Chang & Guenther, 2020) by involving homologous right hemisphere regions. While speculative, those with greater extent of brain anomaly that hampers such compensatory growth may be the risk factors unique to girls who develop persistent stuttering. Further exploration is warranted as the present finding indicates it may be more difficult to identify the females who eventually persist than the males when considering commonly used predictive factors as was explored in the current study.

Clinical Considerations and Implications

Study Sample and External Validity

Study findings indicate that a cumulative risk approach is an appropriate method for screening a child's odds for eventual persistence of stuttering. However, the external validity of our findings, and particularly our cutoff values, has not yet been determined. Our findings are based on our sample of 67 children aged 36–67 months with relatively high speech and language skills and maternal education. Furthermore, our sample's mean time since onset of approximately 12 months and criteria for recovery likely led to a lower recovery rate (i.e., 32%) than samples that include more children closer to stuttering onset at study entry. Validation with new and larger samples of children is needed to determine how well our findings generalize to the wider population of children who stutter. Given our exploratory analysis that indicate predictive factors may yield differential prognostic value for males versus females and findings from Yairi and Ambrose

Figure 5. Stuttering prognostic screener based on the 4-factor model.

Cumulative Risk for Stuttering Persistence

2+ indicators of persistence confers elevated odds to persist

Predictive Factor	Indicating Recovery	Indicating Persistence
Time Since Onset (months)	< 19	≥ 19
Speech-Sound Accuracy (standard score)	> 115	≤ 115
Expressive Language (standard score)	> 106	≤ 106
Stuttering Severity (SSI total score)	< 17	≥ 17
Total Indicators of Persistence (up to 4) Each additional indicator increases child's odds to persist 4.7x		

Note: SSI = Stuttering Severity Instrument
Normative Sample: Sixty-seven 3–5-year-old children with primarily average to above average speech and language skills and an average time since onset of approximately 12 months
Predictive Validity based on sample:
Sensitivity (proportion who persisted accurately identified): 93%
Specificity (proportion who recovered accurately identified): 65%
Positive Predictive Value (proportion predicted to persist, who did): 84%
Negative Predictive Value (proportion predicted to recover, who did): 83%

(2005) indicating that the validity of predictive factors may differ based on a child's time since onset (e.g., stuttering frequency), it is likely the case that alternative cutoff values may be more predictive for children who deviate from our sample (e.g., are closer to stuttering onset or have concomitant speech and/or language disorders). Despite the similarities in predictive factors between our optimal model and the model described in the work of Walsh et al. (2021), further explorations may prove that the selection of predictive factors may vary as well.

Screener

A screener based on our findings, Figure 5, may help speech-language pathologists evaluate a child's prognosis *if* the child is similar to the children included in our sample. Our findings direct clinicians to assess the skills within our optimal model (e.g., speech sound accuracy, stuttering severity, and expressive language) and obtain the child's time since onset to determine how many predictive factors indicate persistence. The factors within this model reflect information that speech-language pathologists report including within comprehensive fluency evaluations for young children (Singer & Kelly, in press), supporting its ecological validity. Children who present with at least two factors indicating persistence are at greater odds for developing persistent stuttering.

Comparison to Alternative Method of Predicting Persistent Developmental Stuttering

The presented cumulative risk approach can be compared to the method for evaluating persistent stuttering explored in Walsh et al. (2021). Walsh et al. (2021) identified profiles of children that may be at greatest risk to persist and encouraged speech-language pathologists to consider which profile best matches the individual child. To identify an individual child's profile, clinicians can consider how the child's scores deviate from the sample's mean. For example, the likelihood of persistence for a child with a moderate stuttering severity based on weighted SLD (their sample's average was 10.1, which is considered moderate) is best determined by the child's score on the Bankson–Bernthal Test of Phonology (BBTOP; Bankson & Bernthal, 1990). Low BBTOP scores (i.e., below their sample's mean of 90.3) increased the child's likelihood of persisting. So, if Child A had a moderate stuttering severity and low speech sound accuracy, they would be at risk to persist based on Walsh et al.'s (2021) findings. Similarly, moderate stuttering severity and a low speech sound score would both be indicative of risk using the cumulative risk method and the child would be identified as being at elevated risk even without considering the last two factors (i.e., time since onset and expressive language). A clinician may elect to not administer an expressive language test for Child A if there are no presenting language concerns as the child's elevated risk has been identified. For Child A, both methods result in the same predicted prognosis, but this will not always be the case. If Child B had moderate stuttering severity and above-average speech sound accuracy (e.g., standard score of 120), they would not be at increased risk to persist based on Walsh et al. (2021). Using the cumulative risk approach, the clinician would be encouraged to also consider the child's time since onset and expressive language skills. If either of these were indicative of risk, the child would be considered to be at elevated risk to persist. The administration of the expressive language test will be particularly important relative to evaluating the child's prognosis if the child has been stuttering for less than 19 months. Last, as opposed to speech-language pathologists considering the figures in Walsh et al. (2021) to identify the child's probability of persistence, the number of predictors indicating risk can be used to calculate a simple *OR* when using the cumulative risk approach (e.g., 9.4 for Child A if no other factors indicating risk were present).

Whereas the method in Walsh et al. (2021) yields higher specificity, we propose that the cumulative risk approach is a viable, simple alternative to the model described in Walsh et al. (2021) given that both approaches yield similarly high sensitivity. When deciding which method to apply to a given case, a speech-language pathologist might consider both the child and which measures were administered. First, the sample in Walsh et al. (2021) included more children with low speech sound skills, so a child with delayed or disordered speech sound skills would be better represented by their study sample. Second, this study utilized the SSI and GFTA, whereas Walsh et al. (2021) utilized weighted SLD and BBTOP as measures of stuttering severity and speech sound accuracy, respectively. Similarly, this study used standardized norm-referenced assessments of language, whereas Walsh et al. (2021) used a nonword repetition task. To date, neither study's findings have been replicated with alternative measures, so it is recommended that the method that utilized similar measures as those used by the speech-language pathologist be selected.

Caveats

Given the retrospective nature of this study, our methods and subsequent findings were limited by the availability of data and there are multiple ways in which future studies might extend our understanding of how well a cumulative risk approach predicts persistent developmental stuttering. This study did not include the temperament/emotion domain; a domain, which despite the limited attention it has received related to stuttering persistence (Ambrose et al., 2015; Singer, Walden, & Jones, 2020), may prove to contribute to predictive validity in the future. Relatedly, we encourage future investigations to include data on child's temperament and reaction toward stuttering to explore negative consequences related to stuttering (beyond persistence) more comprehensively. Other clinical characteristics not available in this study, but found to be predictive of stuttering persistence, include weighted SLD (Walsh et al., 2020) and performance on a nonword repetition task (Spencer & Weber-Fox, 2014). Furthermore, it may be that weighting certain predictive factors, as has been suggested in the literature (e.g., Smith & Weber, 2017; Walsh et al., 2020) and recently explored (Walsh et al., 2021), may improve the predictive validity of a screener.

Recently, Einarsdóttir et al. (2020) identified the importance of considering a child's self-reported status of still stuttering or having recovered when classifying stuttering status. These data had not previously been collected on our study sample, but we believe self-reported status would be important data to collect on future samples. Lastly, inherent to all studies related to stuttering persistence in early childhood, it is possible that children's persistence trajectory changed after the time period in which their data were collected. The longer a child is followed, including after persistence/recovery status has been determined, the chance that a child will change classification will be reduced.

Considering the collaborative nature of this work across research laboratories at differing institutions, the measures employed were not identical in all instances. For example, there were slight differences in the standardized language tests used as well as the approaches for the measurement of stuttering frequency. While these considerations should be noted, our view is that the benefit of combining longitudinal data on developmental stuttering to advance knowledge of predictive outcomes outweighs the current limitations, especially if these initial efforts lead to research laboratories prospectively collaborating on a cohesive set of protocols that will seamlessly enable further empirical study of these critical issues.

Conclusions

This study provides evidence supporting the use of a cumulative risk approach to predict a child's chances of persisting or recovering from stuttering. We found that when considering a child's speech sound accuracy, expressive language skills, stuttering severity, and time since stuttering onset against our data-driven cutoff values, a child's odds of persistent stuttering increase nearly fivefold with the presence of each additional predictive factor indicating persistence. We also found that children with a family history of stuttering persistence and males who persist are likely to present with more risk factors than children without a family history and females who persist, respectively. Whereas it is important that the present results be replicated on new, independent samples of children, our results provide initial support for a practical and simple method of evaluating a child's chances of persisting or recovering from stuttering. These results provide insights into possible future directions that may further help improve predictions of stuttering persistence during early childhood.

Acknowledgments

This work was supported by the National Institute on Deafness and Other Communication Disorders (NIDCD)

Grants R01DC011277 (Chang), R01DC000523 (Conture), R56DC000523 (Walden), and R21DC016723 (Jones); the National Center for Advancing Translational Sciences Grant UL1 TR000445-06 (Bernard); and a Vanderbilt Kennedy Center Hobbs Discovery Grant. This research was also supported by the Wilker-Ellis Stuttering Research Fund at Vanderbilt University Medical Center. The research and content reported herein is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, NIDCD, University of Michigan, Michigan State University, Vanderbilt University, Vanderbilt University Medical Center, the Vanderbilt Kennedy Center, or the generous donors that supported this work. The authors also extend sincere appreciation to the young children and their caregivers who participated in this study, individuals without whose cooperation this project would not have been possible to conduct.

References

Ambrose, N. G., Yairi, E., & Cox, N. (1993). Genetic aspects of early childhood stuttering. *Journal of Speech and Hearing Research, 36*(4), 701–706. https://doi.org/10.1044/jshr.3604.701

Ambrose, N. G., Yairi, E., Loucks, T. M., Seery, C. H., & Throneburg, R. (2015). Relation of motor, linguistic and temperament factors in epidemiologic subtypes of persistent and recovered stuttering: Initial findings. *Journal of Fluency Disorders, 45,* 12–26. https://doi.org/10.1016/j.jfludis.2015.05.004

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.

Bankson, N., & Bernthal, J. E. (1990). *Bankson-Bernthal Test of Phonology*. Pro-Ed.

Blood, G. W., & Blood, I. M. (2004). Bullying in adolescents who stutter: Communicative competence and self-esteem. *Contemporary Issues in Communication Science and Disorders, 31*(Spring), 69–79. https://doi.org/10.1044/cicsd_31_S_69

Carter, J. V., Pan, J., Rai, S. N., & Galandiuk, S. (2016). ROCing along: Evaluation and interpretation of receiver operating characteristic curves. *Surgery, 159*(6), 1638–1645. https://doi.org/10.1016/j.surg.2015.12.029

Chang, S. E., & Guenther, F. H. (2020). Involvement of the cortico-basal ganglia-thalamocortical loop in developmental stuttering. *Frontiers in Psychology, 10,* 3088. https://doi.org/10.3389/fpsyg.2019.03088

Choi, D., Conture, E. G., Tumanova, V., Clark, C. E., Walden, T. A., & Jones, R. M. (2018). Young children's family history of stuttering and their articulation, language and attentional abilities: An exploratory study. *Journal of Communication Disorders, 71,* 22–36. https://doi.org/10.1016/j.jcomdis.2017.11.002

Chow, H. M., & Chang, S.-E. (2017). White matter developmental trajectories associated with persistence and recovery of childhood stuttering. *Human Brain Mapping, 38*(7), 3345–3359. https://doi.org/10.1002/hbm.23590

Clark, C. E., Tumanova, V., & Choi, D. (2017). Evidence-based multifactorial assessment of preschool-age children who stutter. *Perspectives of the ASHA Special Interest Groups, 2*(4), 4–27. https://doi.org/10.1044/persp2.SIG4.4

Conture, E. G., & Walden, T. A. (2012). Dual diathesis-stressor model of stuttering. *Theoretical Issues of Fluency Disorders, 41*(4), 94–127.

Donaher, J., & Richels, C. (2012). Traits of attention deficit/hyperactivity disorder in school-age children who stutter. *Journal of Fluency Disorders, 37*(4), 242–252. https://doi.org/10.1016/j.jfludis.2012.08.002

Dunn, L. M., & Dunn, D. M. (2007). *The Peabody Picture Vocabulary Test–Fourth Edition*. NCS Pearson, Inc.

Einarsdóttir, J. T., Crowe, K., Kristinsson, S. H., & Másdóttir, T. (2020). The recovery rate of early stuttering. *Journal of Fluency Disorders, 64,* 105764. https://doi.org/10.1016/j.jfludis.2020.105764

Evans, G. W., Li, D., & Whipple, S. S. (2013). Cumulative risk and child development. *Psychological Bulletin, 139*(6), 1342–1396. https://doi.org/10.1037/a0031808

Fluharty, N. (2001). *Fluharty Preschool Speech and Language Screening Test–Second Edition (Fluharty-2)*. Pro-Ed.

Garnett, E. O., Chow, H. M., Nieto-Castañón, A., Tourville, J. A., Guenther, F. H., & Chang, S.-E. (2018). Anomalous morphology in left hemisphere motor and premotor cortex of children who stutter. *Brain, 141*(9), 2670–2684. https://doi.org/10.1093/brain/awy199

Gillam, R. B., Logan, K. J., & Pearson, N. A. (2009). *Test of Childhood Stuttering (TOCS)*. Pro-Ed.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation–Second Edition (GFTA-2)*. AGS.

Guitar, B. (2019). *Stuttering: An integrated approach to its nature and treatment* (5th ed.). Wolters Kluwer.

Guitar, B., & Conture, E. G. (2006). *The child who stutters: To the pediatrician*. Stuttering Foundation of America.

Guttormsen, L. S., Kefalianos, E., & Næss, K.-A. B. (2015). Communication attitudes in children who stutter: A meta-analytic review. *Journal of Fluency Disorders, 46,* 1–14. https://doi.org/10.1016/j.jfludis.2015.08.001

Hayiou-Thomas, M. E., Smith-Woolley, E., & Dale, P. S. (2021). Breadth versus depth: Cumulative risk model and continuous measure prediction of poor language and reading outcomes at 12. *Developmental Science, 24*(1), e12998. https://doi.org/10.1111/desc.12998

Hollingshead, A. B. (1975). *Four factor index of social status* [Unpublished manuscript]. Yale University.

Hresko, W. P., Reid, D. K., & Hammill, D. D. (1991). *Test of Early Language Development–Second Edition (TELD-2)*. Pro-Ed.

Ingalhalikar, M., Smith, A., Parker, D., Satterthwaite, T. D., Elliott, M. A., Ruparel, K., Hakonarson, H., Gur, R. E., Gur, R. C., & Verma, R. (2014). Sex differences in the structural connectome of the human brain. *Proceedings of the National Academy of Sciences, 111*(2), 823–828. https://doi.org/10.1073/pnas.1316909110

Kang, C., & Drayna, D. (2011). Genetics of speech and language disorders. *Annual Review of Genomics and Human Genetics, 12,* 145–164. https://doi.org/10.1146/annurev-genom-090810-183119

Kefalianos, E., Onslow, M., Ukoumunne, O., Block, S., & Reilly, S. (2014). Stuttering, temperament, and anxiety: Data from a community cohort ages 2–4 years. *Journal of Speech, Language, and Hearing Research, 57*(4), 1314–1322. https://doi.org/10.1044/2014_JSLHR-S-13-0069

Kelman, E., & Nicholas, A. (2008). *Practical intervention for early childhood stammering: Palin PCI Approach*. Speechmark Publishing Ltd.

Kidd, K. K. (1984). Stuttering as a genetic disorder. In R. F. Curlee & W. H. Perkins (Eds.), *Nature and treatment of stutterings: New direction* (pp. 149–169). College-Hill.

Kidd, K. K., Heimbuch, R. C., & Records, M. A. (1981). Vertical transmission of susceptibility to stuttering with sex-modified expression. *Proceedings of the National Academy of Sciences, 78*(1), 606–610. https://doi.org/10.1073/pnas.78.1.606

Klein, J. F., & Hood, S. B. (2004). The impact of stuttering on employment opportunities and job performance. *Journal of Fluency Disorders, 29*(4), 255–273. https://doi.org/10.1016/j.jfludis.2004.08.001

Kraft, S. J., & Yairi, E. (2012). Genetic bases of stuttering: The state of the art, 2011. *Folia Phoniatrica et Logopaedica, 64*(1), 34–47. https://doi.org/10.1159/000331073

Kraha, A., Turner, H., Nimon, K., Zientek, L., & Henson, R. (2012). Tools to support interpreting multiple regression in the face of multicollinearity. *Frontiers in Psychology, 3,* 44. https://doi.org/10.3389/fpsyg.2012.00044

Langevin, M., Packman, A., & Onslow, M. (2010). Parent perceptions of the impact of stuttering on their preschoolers and themselves. *Journal of Communication Disorders, 43*(5), 407–423. https://doi.org/10.1016/j.jcomdis.2010.05.003

Logan, K. J. (2022). *Fluency disorders: Stuttering, cluttering, and related fluency problems* (2nd ed.). Plural.

Lopez-Raton, M., & Rodriguez-Alvarez, M. X. (2019). *OptimalCutpoints: Computing optimal cutpoints in diagnostic tests* (Version 1.1-4) [Computer software]. https://cran.r-project.org/web/packages/OptimalCutpoints/index.html

Lucio, R., Hunt, E., & Bornovalova, M. (2012). Identifying the necessary and sufficient number of risk factors for predicting academic failure. *Developmental Psychology, 48*(2), 422–428. http://doi.org/10.1037/a0025939

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment,comparison and testing of statistical models. *Journal of Open Source Software, 6*(60), 3139. https://doi.org/10.21105/joss.03139

Lumley, T. (2020). *leaps: Regression subset selection* (Version 3.1) [Computer software]. R package version.

Mallows, C. L. (1973). Some comments on C p. *Technometrics, 15*(4), 661–675. https://doi.org/10.1080/00401706.1973.10489103

Nippold, M. A. (2004). Phonological and language disorders in children who stutter: Impact on treatment recommendations. *Clinical Linguistics & Phonetics, 18*(2), 145–149. https://doi.org/10.1080/02699200310001659070

O'Brian, S., Jones, M., Lincoln, M., Harrison, E., Packman, A., Menzies, R., & Onslow, M. (2013). Validity of real-time measures of stuttering frequency. *Speech, Language and Hearing, 16*(2), 107–116. https://doi.org/10.1179/2050571X13Z.00000000015

Peterson, R. L., McGrath, L. M., Smith, S. D., & Pennington, B. F. (2007). Neuropsychology and genetics of speech, language, and literacy disorders. *Pediatric Clinics of North America, 54*(3), 543–561. https://doi.org/10.1016/j.pcl.2007.02.009

Plante, E., & Vance, R. (1994). Selection of preschool language Tests: A data-based approach. *Language, Speech, and Hearing Services in Schools, 25*(1), 15–24. https://doi.org/10.1044/0161-1461.2501.15

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Riley, G. D. (1994). *Stuttering Severity Instrument for Children and Adults–Fourth Edition*. Pro-Ed.

Riley, G. D. (2009). *Stuttering Severity Instrument for Children and Adults–Third Edition*. Pro-Ed.

Selin, C. M., Rice, M. L., Girolamo, T., & Wang, C. J. (2019). Speech-language pathologists' clinical decision making for children with specific language impairment. *Language, Speech, and Hearing Services in Schools, 50*(2), 283–307. https://doi.org/10.1044/2018_LSHSS-18-0017

Semel, E., Wiig, E., & Secord, W. (2004). *Clinical Evaluation of Language Fundamentals Preschool−2*. Pearson.

Singer, C. M., Hessling, A., Kelly, E. M., Singer, L., & Jones, R. M. (2020). Clinical characteristics associated with stuttering persistence: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 63*(9), 2995–3018. https://doi.org/10.1044/2020_JSLHR-20-00096

Singer, C. M., & Kelly, E. M. (in press). Speech-language pathologists' practices related to evaluating persistence prognosis for children who stutter: A survey study. *Perspectives of the ASHA Special Interest Groups*.

Singer, C. M., Walden, T. A., & Jones, R. M. (2019). Differences in the relation between temperament and vocabulary based on children's stuttering trajectories. *Journal of Communication Disorders, 78,* 57–68. https://doi.org/10.1016/j.jcomdis.2019.01.004

Singer, C. M., Walden, T. A., & Jones, R. M. (2020). Attention, speech-language dissociations and stuttering chronicity. *American Journal of Speech-Language Pathology, 29*(1), 157–167. https://doi.org/10.1044/2019_AJSLP-19-00039

Smith, A., & Weber, C. (2017). How stuttering develops: The multifactorial dynamic pathways theory. *Journal of Speech, Language, and Hearing Research, 60*(9), 2483–2505. https://doi.org/10.1044/2017_JSLHR-S-16-0343

Spencer, C., & Weber-Fox, C. (2014). Preschool speech articulation and nonword repetition abilities may help predict eventual recovery or persistence of stuttering. *Journal of Fluency Disorders, 41,* 32–46. https://doi.org/10.1016/j.jfludis.2014.06.001

Thompson, C. G., Kim, R. S., Aloe, A. M., & Becker, B. J. (2017). Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results. *Basic and Applied Social Psychology, 39*(2), 81–90. https://doi.org/10.1080/01973533.2016.1277529

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research, 39*(6), 1284–1294. https://doi.org/10.1044/jshr.3906.1284

Tumanova, V., Choi, D., Conture, E. G., & Walden, T. A. (2018). Expressed parental concern regarding childhood stuttering and the Test of Childhood Stuttering. *Journal of Communication Disorders, 72,* 86–96. https://doi.org/10.1016/j.jcomdis.2018.01.002

Tumanova, V., Conture, E. G., Lambert, E. W., & Walden, T. A. (2014). Speech disfluencies of preschool-age children who do and do not stutter. *Journal of Communication Disorders, 49,* 25–41. https://doi.org/10.1016/j.jcomdis.2014.01.003

Walsh, B., Bostian, A., Tichenor, S. E., Brown, B., & Weber, C. (2020). Disfluency characteristics of 4- and 5-year-old children who stutter and their relationship to stuttering persistence and recovery. *Journal of Speech, Language, and Hearing Research, 63*(8), 2555–2566. https://doi.org/10.1044/2020_JSLHR-19-00395

Walsh, B., Christ, S., & Weber, C. (2021). Exploring relationships among risk factors for persistence in early childhood stuttering. *Journal of Speech, Language, and Hearing Research, 64*(8), 2909–2927. https://doi.org/10.1044/2021_JSLHR-21-00034

Walsh, B., Usler, E., Bostian, A., Mohan, R., Gerwin, K., Brown, B., Weber, C., & Smith, A. (2018). What are predictors for persistence in childhood stuttering. *Seminars in Speech and Language, 39*(04), 299–312. https://doi.org/10.1055/s-0038-1667159

Williams, K. T. (1997). *Expressive Vocabulary Test–Second Edition*. Pearson, Inc.

Yairi, E., & Ambrose, N. (1992). A longitudinal study of stuttering in children: A preliminary report. *Journal of Speech and*

Hearing Research, 35(4), 755–760. https://doi.org/10.1044/jshr.3504.755
Yairi, E., & Ambrose, N. G. (1999). Early childhood stuttering I: Persistency and recovery rates. *Journal of Speech, Language, and Hearing Research, 42*(5), 1097–1112. https://doi.org/10.1044/jslhr.4205.1097
Yairi, E., & Ambrose, N. G. (2005). *Early childhood stuttering for clinicians by clinicians*. Pro-Ed.
Yairi, E., & Ambrose, N. (2013). Epidemiology of stuttering: 21st century advances. *Journal of Fluency Disorders, 38*(2), 66–87. https://doi.org/10.1016/j.jfludis.2012.11.002
Yairi, E., Ambrose, N., & Cox, N. (1996). Genetics of stuttering. *Journal of Speech and Hearing Research, 39*(4), 771–784. https://doi.org/10.1044/jshr.3904.771
Yairi, E., Ambrose, N. G., Paden, E. P., & Throneburg, R. N. (1996). Predictive factors of persistence and recovery: Pathways of childhood stuttering. *Journal of Communication Disorders, 29*(1), 51–77. https://doi.org/10.1016/0021-9924(95)00051-8
Yaruss, J. S. (1998). Real-time analysis of speech fluency. *American Journal of Speech-Language Pathology, 7*(2), 25–37. https://doi.org/10.1044/1058-0360.0702.25
Yaruss, J. S. (2000). Converting between word and syllable counts in children's conversational speech samples. *Journal of Fluency Disorders, 25*(4), 305–316. https://doi.org/10.1016/S0094-730X(00)00088-7
Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer, 3*(1), 32–35. https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3
Zebrowski, P. M. (1997). Assisting young children who stutter and their families: Defining the role of the speech-language pathologist. *American Journal of Speech-Language Pathology, 6*(2), 19–28. https://doi.org/10.1044/1058-0360.0602.19
Zengin-Bolatkale, H., Conture, E. G., Walden, T. A., & Jones, R. M. (2018). Sympathetic arousal as a marker of chronicity in childhood stuttering. *Developmental Neuropsychology, 43*(2), 135–151. https://doi.org/10.1080/87565641.2018.1432621

Appendix A

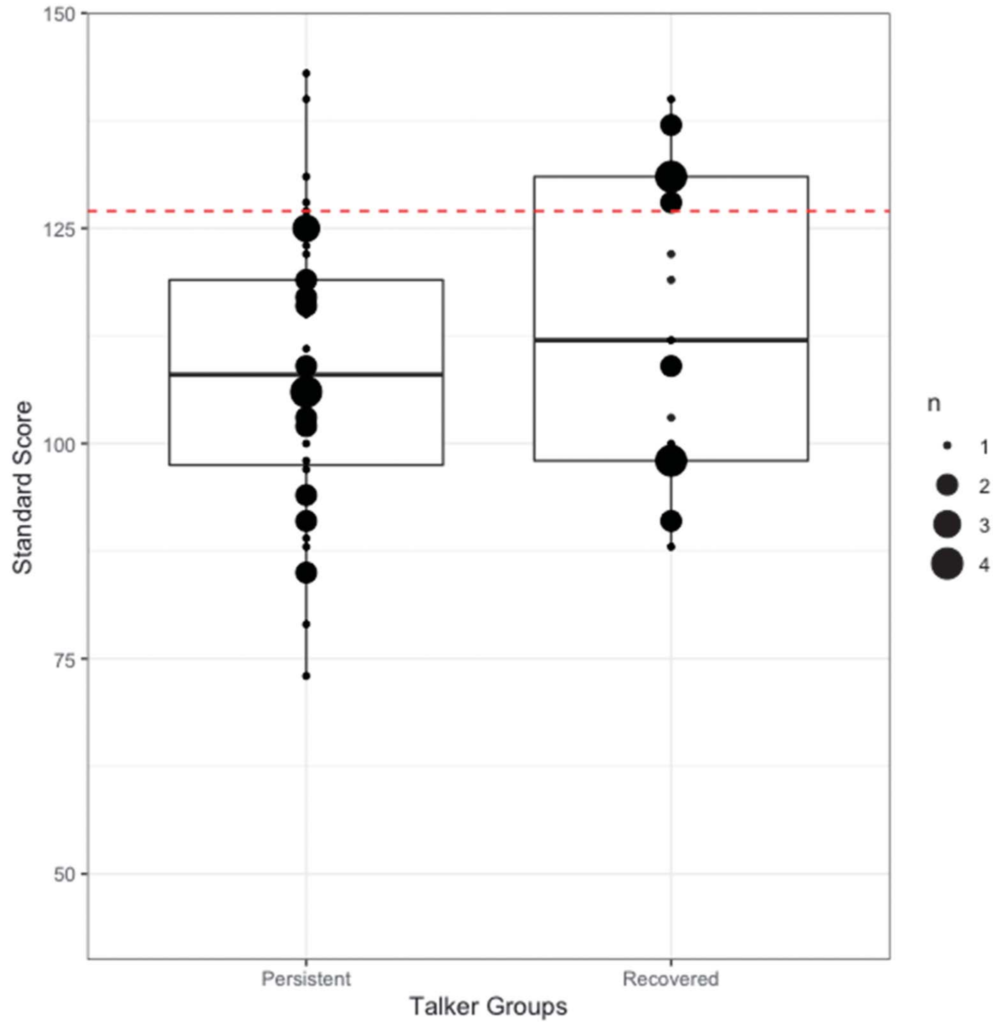
Study Sample Demographics Organized by Original Data Sets

Demographic Variable	Michigan State University	Vanderbilt University Medical Center	Comparison	*p*
Sex	21 M; 13 F	27 M; 6 F	χ^2 (67) = 3.32	.069
Age (months)	50.3 (8.4)	44.5 (6.5)	t(65) = 3.17	**.002**
Maternal education	6.2 (0.8)	6.3 (0.8)	t(65) = −0.67	.503
Age of onset (months)	33.8 (9.2)	32.0 (9.1)	t(65) = 0.82	.417
Stuttering frequency	5.9 (3.0)	7.4 (4.2)	t(65) = −1.72	.090
Time since onset (months)	16.5 (9.3)	12.6 (8.5)	t(65) = 1.83	.072
PPVT (standard score)	116.6 (13.8)	116.8 (12.5)	t(53) = −0.06	.956
EVT (standard score)	113.2 (10.7)	116.9 (11.2)	t(53) = −1.23	.230

Note. Maternal education based on Hollingshead (1975). Stuttering frequency based on percent syllables stuttered. M = male; F = female; PPVT = Peabody Picture Vocabulary Test (Dunn & Dunn, 2007); EVT = Expressive Vocabulary Test (Williams, 1997).

Appendix B

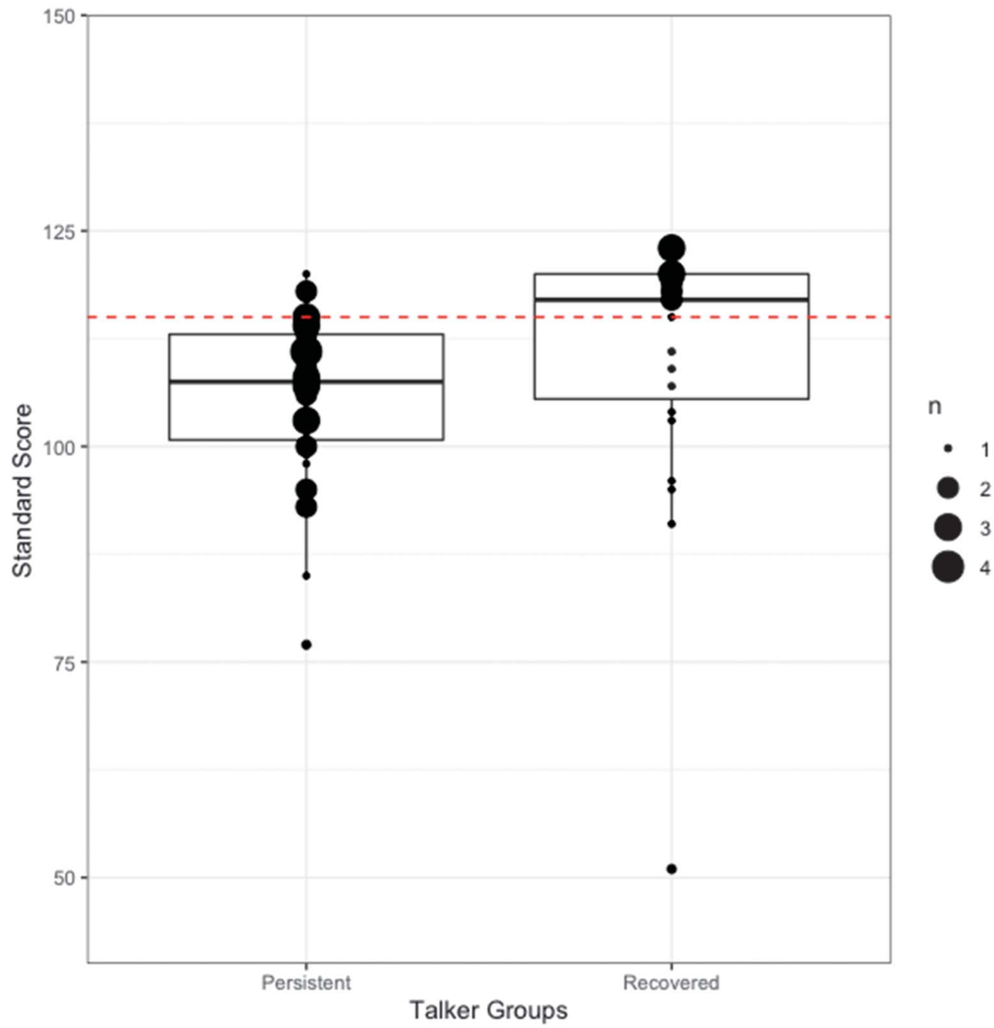
Speech Sound Accuracy Scores by Talker Group



Note. Dashed red line indicates cutoff value based on ROC method.

Appendix C

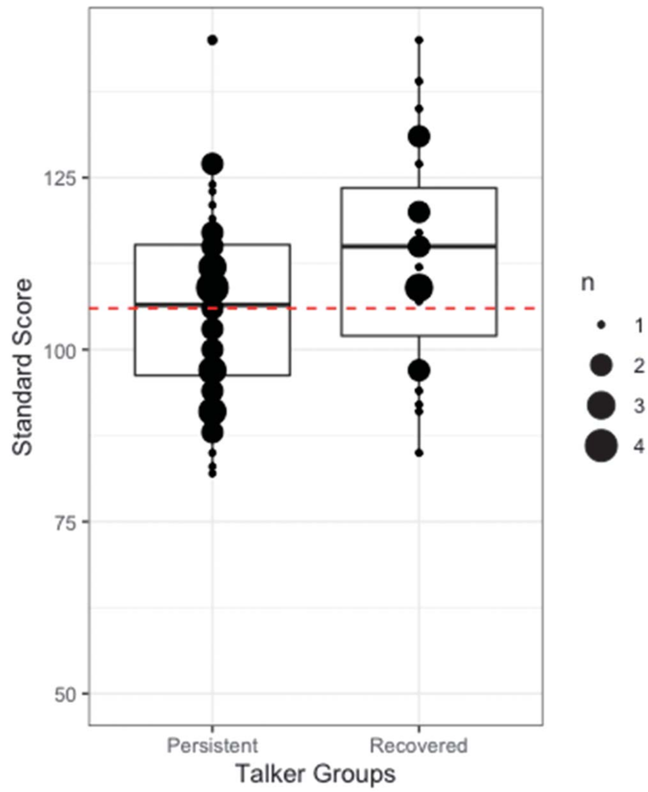
Receptive Language Scores by Talker Group



Note. Dashed red line indicates cutoff value based on ROC method.

Appendix D

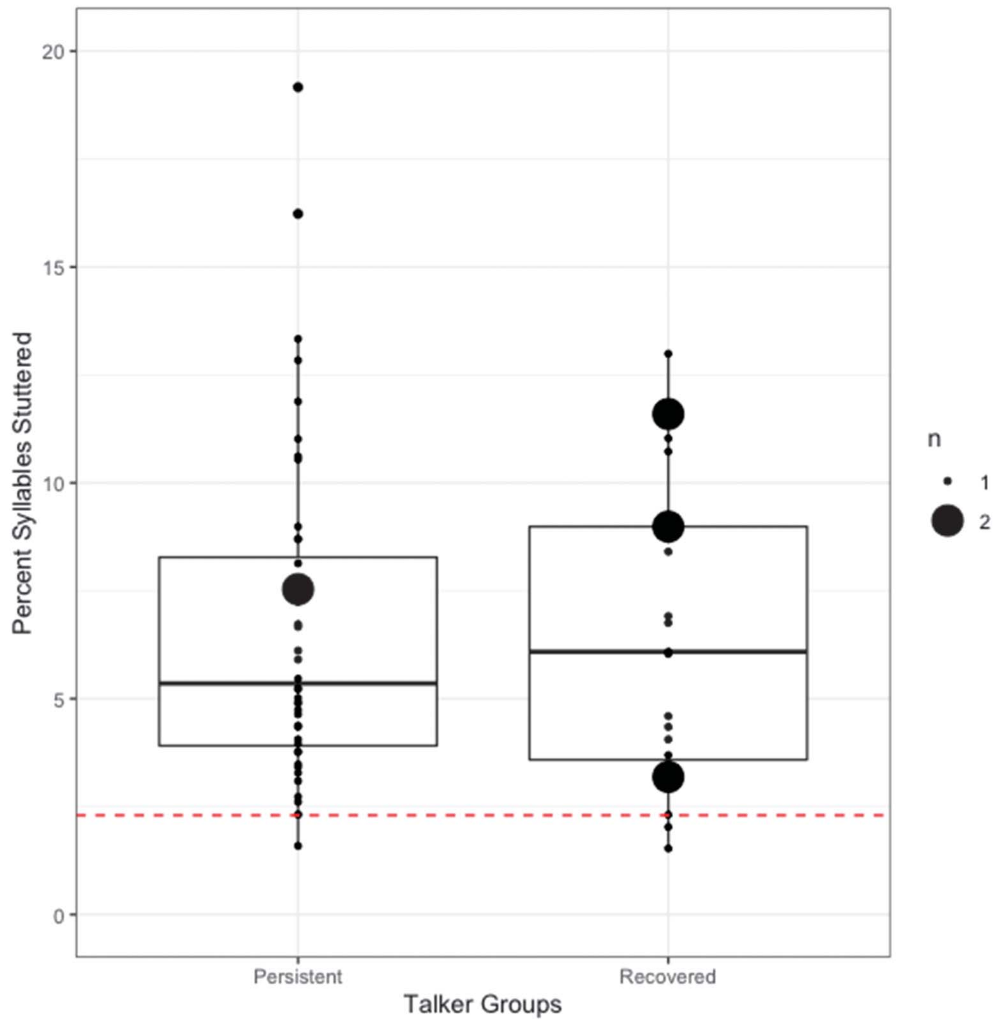
Expressive Language Scores by Talker Group



Note. Dashed red line indicates cutoff value based on ROC method.

Appendix E

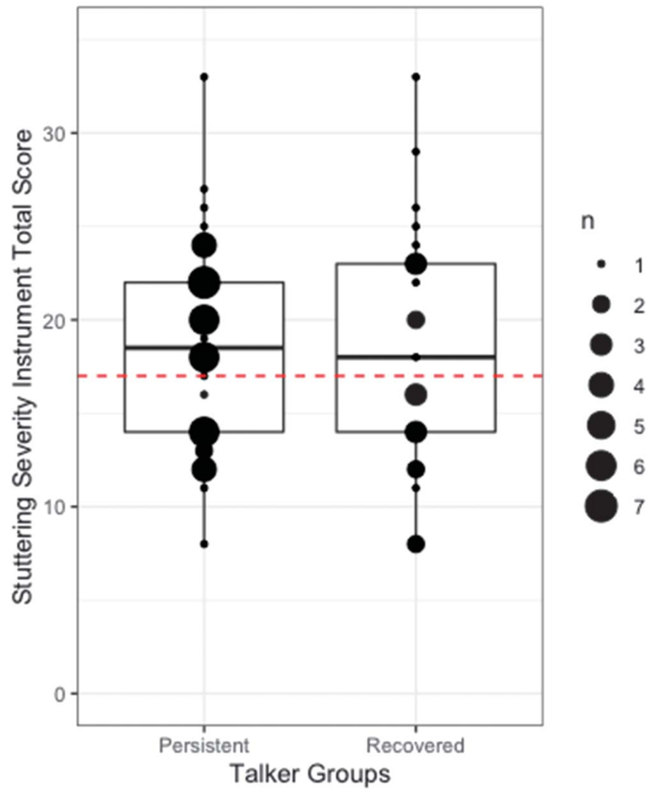
Stuttering Frequency by Talker Group



Note. Dashed red line indicates cutoff value based on ROC method.

Appendix F

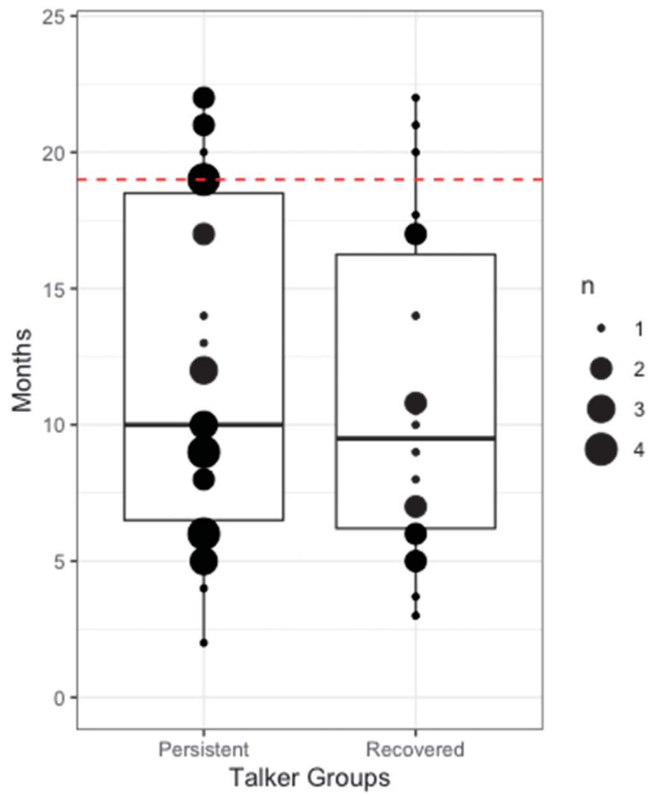
Stuttering Severity Scores by Talker Group



Note. Dashed red line indicates cutoff value based on ROC method.

Appendix G

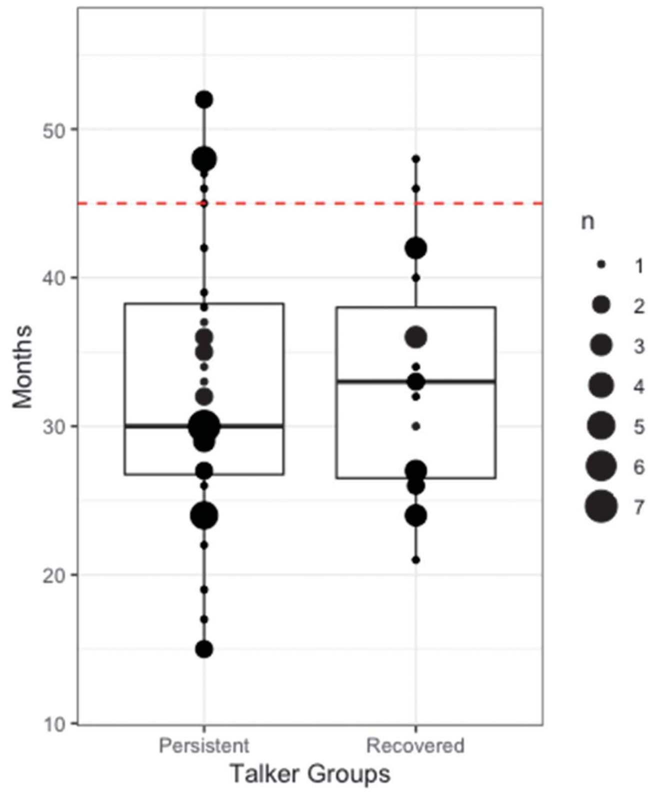
Time Since Onset by Talker Group



Note. Dashed red line indicates cutoff value based on ROC method.

Appendix H

Age at Onset by Talker Group



Note. Dashed red line indicates cutoff value based on ROC method.