JSLHR

Clinical Focus

Exploring Relationships Among Risk Factors for Persistence in Early Childhood Stuttering

Bridget Walsh,[a] Sharon Christ,[b] and Christine Weber[c]

Purpose: The purpose of this study is to investigate how epidemiological and clinical factors collectively predict whether a preschooler who is stuttering will persist or recover and to provide guidance on how clinicians can use these factors to evaluate a child's risk for stuttering persistence.
Method: We collected epidemiological and clinical measures from 52 preschoolers (*M* = 54.4 months, *SD* = 6.7 months; 38 boys and 14 girls) diagnosed as stuttering. We then followed these children longitudinally to document whether they eventually recovered or persisted in stuttering. Risk factors found to be significantly associated with stuttering persistence were used to build single and multiple variable predictive statistical models. Finally, we assessed each model's prediction capabilities by recording how accurate a model was in predicting a child's stuttering outcome—persisting or recovered.
Results: We found that a positive family history of stuttering, poorer performance on a standardized articulation/phonological assessment, higher frequency of stuttering-like disfluencies during spontaneous speech, and lower accuracy on a nonword repetition task were all significantly associated with an increased probability of persistence. The interaction between family history of stuttering and nonword repetition performance was also significant. The full multiple regression model incorporating all these risk factors resulted in the best fitting model with the highest predictive accuracy and lowest error rate.
Conclusions: For the first time, we show how multiple risk factors collectively predict the probability of stuttering persistence in 3- to 5-year-old preschool children who stutter. Using the full combination of risk factors to assess preschoolers who stutter yielded more accurate predictions of persistence compared to sparser models. A better understanding of the factors that underlie stuttering persistence will yield insight into the underpinnings of chronic stuttering and will help identify etiological targets for novel treatment approaches.

Childhood onset fluency disorder, also referred to as childhood stuttering, is a neurodevelopmental disorder that affects 5%–11% of preschool-age children (Bloodstein & Bernstein Ratner, 2008; Reilly et al., 2013; Yairi & Ambrose, 2005). Given the relatively high incidence of stuttering, a significant concern is how to diagnostically differentiate children more likely to recover from children at greater risk for persisting and developing a chronic speech disorder. Stuttering emerges in young children, on average, around 33 months (Yairi & Ambrose, 2005). The prognosis for recovery is favorable with 75%–80% of children recovering from stuttering within 15 months of onset (Yairi & Ambrose, 1999; Yairi et al., 1993). Consequently, parents may be counseled to delay therapy to allow time for natural remission. Recovery rates decline, however, to 50%–60% by the time a child reaches age 5 years (Walsh et al., 2018, 2020; Yairi & Ambrose, 2005). Delaying therapy for children at greater risk for persisting allows atypical neural speech motor networks to form (Chang & Zhu, 2013; Garnett et al., 2018; Hosseini et al., 2018; Walsh et al., 2017), resulting in a challenging speech disorder to treat. Preschoolers who continue to stutter are at risk for developing negative communication attitudes and psychological distress that may adversely affect their lives and reduce the effectiveness of therapy (De Nil & Brutten, 1990; Tran et al., 2011; Vanryckeghem et al., 2005; Yaruss, 2010). Beyond age 7 years, children who are persisting are at significant risk for chronic stuttering with potentially negative consequences for psychosocial development and academic

[a]Department of Communicative Sciences and Disorders, Michigan State University, East Lansing
[b]Department of Human Development and Family Studies, Purdue University, West Lafayette, IN
[c]Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, IN
Correspondence to Bridget Walsh: walshb16@msu.edu

Editor-in-Chief: Bharath Chandrasekaran
Editor: Julie D. Anderson
Received January 21, 2021
Revision received March 13, 2021
Accepted April 7, 2021
https://doi.org/10.1044/2021_JSLHR-21-00034

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

and vocational achievement (Blumgart et al., 2010; Craig et al., 2009; Klein & Hood, 2004; O'Brian et al., 2011).

To date, research pinpoints epidemiological and clinical risk factors associated with stuttering persistence (for recent meta-analysis, see Singer et al., 2020). However, it is unknown how these factors may be utilized for understanding the probability for persistence in children who begin to stutter. Furthermore, an understanding of the relationships among these critical factors when taken together may provide more powerful means for predicting persistence. Evidence suggests that the developmental pathways for stuttering are affected by multiple, dynamic, and interacting factors (Smith & Weber, 2017). Discovering a combination of risk factors and their relationships to help more accurately predict whether a preschool child's stuttering will persist is a vital goal. This knowledge would advance understanding of the underpinnings of chronic stuttering, allow clinicians to prioritize resources for those children in immediate need of intervention, and may help identify targets for therapies. Over the past 15 years, our longitudinal research has explored physiological, behavioral, and clinical factors associated with stuttering persistence and recovery in preschool children (for a review, see Walsh et al., 2018). In this prospective study, we focus on a subset of risk factors available to clinicians that were associated with stuttering persistence and recovery in ours and other's earlier work (Spencer & Weber-Fox, 2014; Walsh et al., 2020; Yairi & Ambrose, 2005). We examined whether epidemiological and clinical factors recorded when children who stutter (CWS) entered the study, between the ages of 3 and 5 years collectively confer an increased risk for stuttering persistence. This research is essential to better understand how speech-language pathologists (SLPs) can use prognostic factors of stuttering persistence to make consequential decisions about a child's risk for developing chronic stuttering. In this study, we will examine, for the first time, how epidemiological and clinical risk factors collectively predict stuttering outcomes—persistence or recovery—in preschool CWS.

Predictive Factors of Stuttering Persistence

Epidemiological Factors

Epidemiological factors are often used to evaluate a child's risk of stuttering persistence. These factors were established, in part, through longitudinal documentation of the onset and development of stuttering in preschool children (Reilly et al., 2009; Yairi & Ambrose, 2005). Yairi and Ambrose (2005) refer to a set of these predictors as "primary factors" or risk factors that are strongly associated with stuttering persistence. One such prognostic factor is family history. The familial incidence of stuttering is well established (Bloodstein & Bernstein Ratner, 2008), although the precise genetic transmission of stuttering is unclear. There is a considerable range in the frequency of stuttering among relatives of adults who stutter from 10.3% to as high as 84% (Janssen et al., 1996; Viswanath et al., 2004). In a large-scale prospective community-cohort study, the Early Language in Victoria Study, approximately 52% of 137 preschoolers who had begun to stutter had a positive family history of stuttering (Reilly et al., 2009). Approximately 66% of a sample of 86 preschool CWS identified a first-, second-, or third-degree relative to have ever stuttered (Yairi & Ambrose, 1992a). Yairi et al. (1996) noted that it is also important to consider whether that family member recovered or persisted from stuttering. They examined stuttering pedigrees from 12 children who persisted in stuttering and 20 children who recovered from stuttering and found that CWS were more likely to persist or recover from stuttering if they had a relative who persisted or recovered themselves. Specifically, children who persisted with a positive family history of stuttering reported that 5.5% of their relatives stuttered and persisted, while 1% of their relatives recovered. Children who recovered from stuttering reported that 1.5% of their relatives stuttered and persisted, while 4.3% of their relatives recovered.

Another primary risk factor for persistence is sex. The male-to-female stuttering ratio for adults is estimated to be 4:1 or larger (Bloodstein & Bernstein Ratner, 2008). Yairi and Ambrose (1992a) reported a 2:1 male-to-female ratio in a group of 87 preschoolers ranging in age from 23 to 75 months. In our cohort of 4- and 5-year-old CWS, we found boys outnumbered girls by a ratio of nearly 3:1 (Walsh et al., 2018). The sex ratio thus increases with age, indicating that preschool-age girls are more likely to recover from stuttering than their male peers (Dworzynski et al., 2007; Reilly et al., 2009; Yairi & Ambrose, 2005).

The age at stuttering onset is also considered a primary risk factor for persistence. Yairi and Ambrose (2005) reported that the average age at stuttering onset in preschoolers is 33 months, with 60% of onsets occurring prior to children's third birthday, and nearly all, or 95%, of stuttering onsets occurring by age 48 months. Yairi and Ambrose (2005) found that children who eventually persisted (CWS-ePer) began stuttering approximately 3.5 months later than children who would eventually recover (CWS-eRec), although this was not a statistically significant difference. Our longitudinal data also indicated that CWS-ePer (n = 18; M = 36.3 months) began stuttering approximately 2 months later than the group of CWS-eRec (n = 29; M = 34.3 months). This difference was also not statistically significant; the two groups largely overlapped (Walsh et al., 2020). The recent meta-analysis by Singer et al. (2020) reported later ages at stuttering onset for children who would eventually persist.

A final primary epidemiological risk factor to consider is the duration that a child has been stuttering. In 75%–80% of cases, stuttering resolves within 15 months of onset (Yairi & Ambrose, 2005). It follows then that the chance of recovery decreases as a child continues to stutter past 15 months. In our recent study, we noted a lower recovery rate of approximately 62% for a cohort of forty-seven 4- and 5-year-old CWS (Walsh et al., 2020). These children had already been stuttering, on average, 21 months, so our sample did not include children whose stuttering resolved earlier than age 4–5 years. We found statistically similar parent-reported duration of stuttering between CWS-ePer (M = 22.1 months) and CWS-eRec

(*M* = 20.4 months) and concluded that duration of stuttering in 4- and 5-year-old CWS may not be as useful an indicator of risk as it is for children at other ages.

Severity of Stuttering-Like Disfluencies

Preschoolers commonly produce "typical" disfluencies as a natural part of speech acquisition. Typical disfluencies are produced by most speakers and include revisions, multisyllabic word or phrase repetitions, hesitations/pauses, and interjections. Yairi and Ambrose (1999) confirmed that the frequency of typical disfluencies did not differentiate groups of CWS and children who do not stutter (CWNS). We and others also note that the frequency of typical disfluencies is statistically similar between groups of CWS-ePer and CWS-eRec (Singer et al., 2020; Walsh et al., 2020). On the other hand, stuttering-like disfluencies (SLDs)—sound prolongations, blocks, sound/syllable repetitions, and monosyllabic whole-word repetitions—not only differentiate CWS from those who do not stutter but also distinguish CWS-ePer and CWS-eRec (Walsh et al., 2020). In this study, we calculated a composite weighted stuttering-like disfluency (WSLD) index (Ambrose & Yairi, 1999), a severity measure that considers the frequency, type, and number of repetitions of stuttering-like disfluencies (SLD), produced during spontaneous speech (Walsh et al., 2020). Participants were divided into groups of children who would eventually persist (CWS-ePer; *n* = 18) or recover (CWS-eRec; *n* = 29) based on their diagnosis the final year of participation in the longitudinal study. We found that children with higher WSLD scores at ages 4–5 years were significantly more likely to persist. This finding is notable, as previous longitudinal studies of younger CWS did not find a significant relationship between SLD frequency and eventual stuttering outcome (Throneburg & Yairi, 2001; Yairi & Ambrose, 1992b). These studies revealed that although CWS-eRec showed a steeper rate of decline in SLD over time compared to CWS-Per, the difference between the two groups did not reach statistical significance in studies with CWS aged 4 years and younger. We found that by ages 4–5 years, the WSLD does help identify children who are at greater risk for stuttering persistence (Walsh et al., 2020). Diverging pathways to recovery and persistence are already underway in these older preschoolers and are reflected in the average differences in frequency and type of SLD.

Linguistic Risk Factors

Stuttering emerges during a period of marked advancement in children's linguistic abilities (Reilly et al., 2009). There have been a number of studies examining potential relationships between stuttering and language development in groups of preschool CWS and age-matched typically fluent peers (Bauman et al., 2012; Bloodstein & Bernstein Ratner, 2008; Ntourou et al., 2011; Singer et al., 2019; Wagovich et al., 2009; Yairi & Ambrose, 2005). Previous reports also cite concomitant speech and language deficits as risk factors for stuttering persistence (Singer et al., 2020; Yairi & Ambrose, 2005; Yaruss et al., 1998). We were interested to learn whether language proficiency, as measured by standardized and other assessments, differentiated groups of children whose stuttering eventually persisted or resolved (Leech et al., 2017; Spencer & Weber-Fox, 2014). Leech et al. (2017) assessed the development of productive syntax and vocabulary diversity in 50 CWS by analyzing spoken language samples collected over the course of several years. Results revealed that children with steeper productive syntactic growth were more likely to recover from stuttering (Leech et al., 2017). In a prospective study, forty 3- to 5-year-old CWS completed a clinical battery comprising assessments of expressive language, the Structured Photographic Expressive Language Test–Third Edition (Dawson et al., 2003), receptive language, the Test of Auditory Comprehension of Language–Third Edition (Carrow-Woolfolk, 1999), phonological and articulation abilities, the Bankson–Bernthal Test of Phonology (BBTOP; Bankson & Bernthal, 1990), Auditory and Word Memory subtests of the Test of Auditory Perceptual Skills–Revised (Gardner, 1985), and the nonword repetition test (NRT; Dollaghan & Campbell, 1998) to determine whether performance on these measures helped identify children at greater risk for stuttering persistence (Spencer & Weber-Fox, 2014). After the initial visit when the testing battery was completed, we evaluated children's speech each year of the longitudinal study to form groups of CWS-eRec (*n* = 21) and CWS-ePer (*n* = 19). Binary logistic regression analysis confirmed that lower scores on the consonant inventory of the BBTOP (Bankson & Bernthal, 1990) and on a nonword repetition accuracy test (Dollaghan & Campbell, 1998) each significantly predicted eventual stuttering persistence. Scores on receptive and expressive language assessments and verbal working memory measures were not significantly different between the two groups of children (Spencer & Weber-Fox, 2014). We concluded that consonant and nonword production accuracy may reveal key production abilities that may help predict the course of stuttering in some CWS.

Aim of the Current Study

In the multifactorial dynamic pathways (MDP) theory of stuttering (Smith & Weber, 2017), we proposed that complex, nonlinear interactions among multiple factors contribute to the development of stuttering. This theory focuses experimental attention on early childhood, the developmental window of the preschool years in which stuttering emerges then follows a developmental trajectory toward recovery or persistence. In earlier studies from our research group, we used physiological, behavioral, and clinical measures to discover whether patterns in performance were predictive of stuttering recovery or persistence (Hosseini et al., 2018; Leech et al., 2017; see the review in Walsh et al., 2018, 2020). Along with providing additional support for primary risk factors (e.g., sex and family history), our research revealed additional factors that helped establish a neurophysiological bases for stuttering and were associated with persistence and recovery. For example, event-related potentials revealed processing differences between CWS-ePer and CWS-eRec on receptive language tasks (Kreidler et al., 2017; Mohan & Weber, 2015; Usler & Weber-Fox, 2015). Patterns of brain activity recorded during speech production with functional near-infrared spectroscopy also distinguished children who had either persisted or recovered from stuttering (Hosseini et al., 2018).

Earlier studies from our lab and others have contributed substantially to our understanding of how stuttering develops in preschool children by revealing *individual* factors associated with stuttering persistence and recovery. The overall aim of the current study is to advance this research by examining the relationships among *multiple* risk factors motivated by earlier work in the same group of preschoolers who stutter to eventual outcomes of persistence or recovery. Our objectives are twofold. First, we will build statistical models that include combinations of clinically accessible factors revealed through ours and other's earlier studies (i.e., epidemiological, linguistic, and stuttering behaviors) to assess their ability to predict persistence. Second, we will evaluate the diagnostic accuracy of each model that significantly predicts stuttering persistence by calculating sensitivity, specificity, and predictive value measures along with error rates. The second objective adds translational value to our predictive models by establishing how accurate they are at diagnosing outcome—persistence or recovery. Considering multiple factors when assessing a child's overall risk for persistence is clearly important given that children will manifest different profiles of risk factors. For example, not all children whose stuttering persisted in our studies have a positive family history of stuttering, nor did all children who recovered from stuttering complete the nonword repetition task with high accuracy. Given that a unique set of factors contributes to the onset and development of stuttering in each child who is stuttering (Smith & Weber, 2017), we hypothesize that a combination of risk factors will yield greater precision predicting the probability of stuttering persistence compared to the predictive capabilities of risk factors in isolation.

Method

Participants

Fifty-two children (38 boys and 14 girls) participated in the longitudinal study. Note that the majority of these children's data have been reported in our earlier studies (i.e., Spencer & Weber-Fox, 2014; Walsh et al., 2018, 2020). Data were collected at Purdue University and at The University of Iowa using identical assessment and experimental protocols under the approval of the institutional review boards of both universities. At the onset of the longitudinal study, participants eligible for this study were under the age of 6 years (between 41 and 68 months, M = 54.4 months, SD = 6.7 months). This study was part of a larger project on the development of stuttering, and children under age 4 years generally could not be expected to complete the physiological (e.g., kinematic, electroencephalographic) protocols. All participants were native American English speakers; passed a standard hearing screening at 500, 1000, 2000, 4000, and 6000 Hz at 20 dB HL; had normal or corrected-to-normal vision; and had no history of neurological disorders or major illnesses per parent/legal guardian report. The participants scored within normal limits on assessments of nonverbal intelligence, the Columbia Mental Maturity Scale (Burgemeister et al., 1972), and social development, the Childhood Autism Rating Scale (Schopler et al., 1988).

Stuttering Diagnosis

The children were diagnosed as stuttering based on the following criteria developed by Yairi and Ambrose (1999): (a) The parent(s) considered the child to be stuttering, (b) the project SLP with expertise in childhood stuttering considered the child to be stuttering, (c) the SLP rated the child a "2" or higher on an 8-point scale (0–1 = *normal*; 2–3 = *mild stuttering*; 4–5 = *moderate stuttering*; 6–7 = *severe stuttering*). This rating was based upon the type (i.e., sound prolongations, blocks, sound/syllable repetitions, monosyllabic whole-word repetitions), duration, and frequency of SLD along with the presence of secondary characteristics (e.g., body movement, eye blinks, tensing articulatory muscles), and (d) the child exhibited three or more SLD per 100 syllables (3% SLD) collected across two conversational speech samples with the SLP and parent. Details of speech sample collection and analysis are provided in our earlier publication (Walsh et al., 2020). Briefly, we audio- and video-recorded children's play-based interactions with their parent and with the project SLP. The samples were transcribed and typical and SLDs coded using the Systematic Analysis of Language Transcripts program (Miller & Iglesias, 2006) by a trained graduate research assistant and by the project SLP. Discrepancies in transcriptions or coding of disfluencies were resolved by the SLP and research assistant by reviewing the item in question and reaching a consensus.

In 23% of cases, CWS met all criteria except the fourth (i.e., they produced fewer than 3% SLD—on average, 2.5% SLD) during the collection of the speech sample described above, yet parents indicated that the speech sample was not representative of their child's stuttering at home or in other contexts. Fluctuations in stuttering frequency across different situations is a hallmark characteristic of the disorder (Constantino et al., 2016; Yaruss, 1997). In these cases, the SLP considered the child's disfluencies in other contexts such as during standardized speech/language testing or while completing other portions of the experimental protocol when diagnosing stuttering.

Stuttering Classification

Children returned to the lab annually to have their speech reevaluated. After their initial visit, participants were followed, on average, 3.2 years (SD = 1.08 years). Children were considered recovered if they met the following criteria (Yairi & Ambrose, 1999): (a) The child's parents no longer considered their child to be stuttering and rated their stuttering severity to be a 0 or 1 (*normal fluency*) on the 8-point scale, (b) the project SLP no longer considered the child to be stuttering and also rated their stuttering severity to be a 0 or 1 on the 8-point scale, and (c) the child produced less than 3% syllables stuttered during the spontaneous speech samples. If a child did not meet all these criteria, they were considered persisting. We used a child's status from their final year of participation to classify them as persisting or recovered. Of the 52 CWS, 31 children or 59.6% were identified as a child who eventually recovered from stuttering and 21 or 40.4% of children's stuttering persisted (CWS-ePer). Appendix A provides epidemiological information for all CWS-ePer and CWS-eRec including sex, ages at first and last visits to the lab, age of stuttering onset, duration a child had been stuttering at the time of their first visit, and family history. We aimed to collect three or more time points in this longitudinal study to ascertain whether a CWS had recovered or was persisting. We were successful in over 90% of cases. For the five children whose families were unable to return to the lab to participate in the longitudinal study for their third year of participation, we followed up with the family in the later years of the project to obtain an update on a child's speech status. The project SLP documented that two CWS-ePer who were persisting at their second-year visit when they were 62 and 64 months, respectively, were still stuttering when they were 8 years old. We used the Year 2 status for the remaining three children for whom we were not able to get an update (see Appendix A).

Measures

Epidemiological Factors

Age of stuttering onset/duration of stuttering. Parent(s) reported their child's age (in months) of stuttering onset. They were asked to recollect the specific time of year and events surrounding the onset to facilitate recall using the approach described by Yairi and Ambrose (1992a, p. 783). The duration of stuttering was calculated by subtracting a child's current age in months from the parent-reported age of stuttering onset.

Family history. Family history was treated as a binary variable. Parent(s) were asked to report whether there were first through third degree relatives (e.g., siblings, parents, grandparents, aunts and uncles, first cousins, and great-grandparents) who stuttered and whether family members' stuttering persisted or resolved.

Speech/language assessments. We administered a comprehensive speech and language assessment battery to all participants during their initial year of participation in the study (Spencer & Weber-Fox, 2014). The current study focuses on those assessments showing significant group differences between CWS-eRec and CWS-ePer revealed in our earlier reports (Spencer & Weber-Fox, 2014; Walsh et al., 2018). Standardized scores from the BBTOP, Consonant and Phonological Process Inventory (BBTOP-CI, -PPI) subtests (Bankson & Bernthal, 1990) captured participants' consonant articulation and phonological production accuracy. We administered the NRT (Dollaghan & Campbell, 1998) to assess nonword repetition skills. The NRT stimuli are 16 nonwords: four 1-syllable nonwords, four 2-syllable nonwords, four 3-syllable nonwords, and four 4-syllable nonwords. The children were told they would hear a new "alien word" one time and to repeat it as best they could. Recordings of the 16 nonwords were produced by a female native English-speaking adult. We scored phonemes including vowels and consonants for each nonword as correct if they were produced accurately, regardless of whether they were produced fluently. Given that children with speech sound disorders often perform with reduced accuracy on NRT (e.g., Munson et al., 2005), if a child demonstrated a phonological error on the BBTOP, the phoneme was not scored as incorrect on the NRT (Spencer & Weber-Fox, 2014). We included each participant's overall nonword repetition accuracy percentage, or the percentage of phonemes produced correctly across all 16 nonwords out of 96 possible phonemes in the analysis, henceforth NRT score, in the statistical models.

Finally, we calculated a WSLD, a comprehensive index of stuttering severity, for each child using the speech samples described under "Stuttering Diagnosis" (Ambrose & Yairi, 1999). The WSLD is a stuttering severity measure capturing the frequency, type, and extent of SLD into a single score. Note that the WSLD may be computed using FLUCALC, within the CHILDES CLAN software (https://doi.org/10.21415/T5G10R; MacWhinney, 2000; Ratner et al., 1996). It is calculated by adding together part word (PW) and single-syllable (SS) whole word repetitions per 100 syllables of speech and then multiplying this value by the mean number of iterations or repetition units (RUs). Next, this value is added to 2 times the average number of dysrhythmic phonations (DPs) resulting in the equation $[(PW + SS) \times \text{mean RU}) + (2 \times DP) = WSLD]$. DPs receive a weighting because they infrequently occur in early childhood stuttering (Yairi & Ambrose, 2005). A score of ≥ 4.0 is used to diagnose stuttering in children, with 4.00–9.99 denoting mild stuttering, 10.00–29.99 moderate stuttering, and ≥ 30.00 severe stuttering (Ambrose & Yairi, 1999).

Data Analysis

Statistical Model Building

Our analysis approach to delineate risk factors and assess their collective ability to predict stuttering persistence and recovery involved four separate steps outlined in the next two sections. For Step 1, we estimated bivariate logistic regression models to determine which of the predictor factors were associated with stuttering persistence based on group data. Both linear and nonlinear associations were tested where the nonlinear association models included a squared component (quadratic) of the predictor factor. In Step 2, we estimated a second set of models to evaluate possible two-way interactions among the predictor factors. For Step 3, we then built multiple variable logistic regression models using the predictor factors found to be linearly or nonlinearly associated with stuttering persistence in the bivariate models. Any two-way interactions that were found in the interaction models were also included in the multiple variable model building. We initially estimated a full, multiple variable model that included all statistically significant (alpha ≤ .05) predictors from the bivariate (linear and nonlinear) and two-way interaction models. The equation for this full model is given in Equation (1):

$$ln\left(\frac{p}{1-p}\right) = b_0 + b_1\text{FamHist} + b_2\text{WSLD} + b_3\text{WSLD}^2 + b_4\text{NRT}_{\text{Score}} + b_5\text{FamHist}_X\text{NRT}_{\text{Score}} + b_6\text{BBTOP-PPI} \quad (1)$$

where $ln\left(\frac{p}{1-p}\right)$ is the natural log of the odds of stuttering persistence (i.e., the logit transformation). The b_0 is a constant term, b_1 gives the effect of family history of stuttering, b_2 gives the effect of the WSLD score, b_3 gives the effect of WSLD squared, b_4 gives the effect of NRT score, b_5 gives the effect of the interaction between family history and the NRT score, and b_6 gives the effect of the BBTOP-PPI score. Predictions of stuttering persistence using the full model are based on the regression (beta) weights that we get for each factor in the full model equation. The full model considers each factor after controlling for other factors along with significant interactions (i.e., additive and combined effects). We also tested nested submodels that included only those predictor factors meeting statistical significance at an alpha level ≤ .05 in the full multiple variable model relative to the full model using log-likelihood ratio tests. Factors were centered for quadratic and interaction estimation. Stata Version 15.1 was used for all analyses.

Diagnostic Accuracy

Finally, in Step 4a, the model-based predicted probability of stuttering persistence (generated from group data) for the bivariate regression models, multiple variable full model, and multiple variable submodels were then compared with the actual stuttering outcome for each participant. First, we computed sensitivity, specificity, and positive and negative predictive values for the bivariate regression models, multiple variable full model, and multiple variable submodels for a range of predictive probability cutoffs ranging from more to less conservative: 0.30, 0.40, 0.50, and 0.60. For example, a cutoff of .30 for a given model indicates that if that model revealed that a child's risk for persistence was 30% or higher, then that child would be considered a candidate for immediate intervention. Sensitivity is the proportion of people with a condition who are correctly identified as having that condition, and specificity is the proportion of people without a condition correctly identified as not having that condition. In screening contexts, Trevethan (2017) argued the importance of including two additional metrics: positive predictive values (probability that people with a positive screening result have the condition) and negative predictive values (probability that people with a negative screening result do not have the condition) in addition to sensitivity and specificity to provide a more cohesive picture. Finally, in Step 4b, we also calculated error rates, or the total number of children incorrectly predicted to persist (false positives) plus children who were incorrectly predicted to recover (false negatives) out of the total number of participants, at each model-predicted probability cutoff score: 0.30, 0.40, 0.50, and 0.60.

Results

Model results for each independent predictor factor are presented in Table 1 as odds ratios (*OR*s) along with the 95% confidence intervals (CIs) and *p* values for the *OR*s. *OR*s are a common effect size measure for binary outcomes where the effect is the odds of persistence versus recovery for a unit change in the independent variable (i.e., predictor factor). *OR*s over 1 indicate that the odds of persistence versus recovery increased for every unit increase in the predictor factor, while *OR*s less than 1 indicate that the odds of persisting decrease for every unit increase in the predictor factor. *P* values indicate whether the estimated effect is significantly different from 1 (where 1 is no difference in the odds of persistence versus recovery). *OR*s from the models were transformed into predicted probabilities of stuttering persistence for various levels of the predictor factors and presented graphically in Figures 1–3.

Single-Variable Analysis

Sex and Age

Bivariate regression model estimates are presented for each individual predictor factor in Table 1. Out of the initial sample of 14 females, four (28.6%) persisted and 10 (71.4%) recovered. For the group of 38 boys, 17 persisted (44.7%) and 21 recovered (55.3%). Sex was not associated with eventual status in the bivariate model.

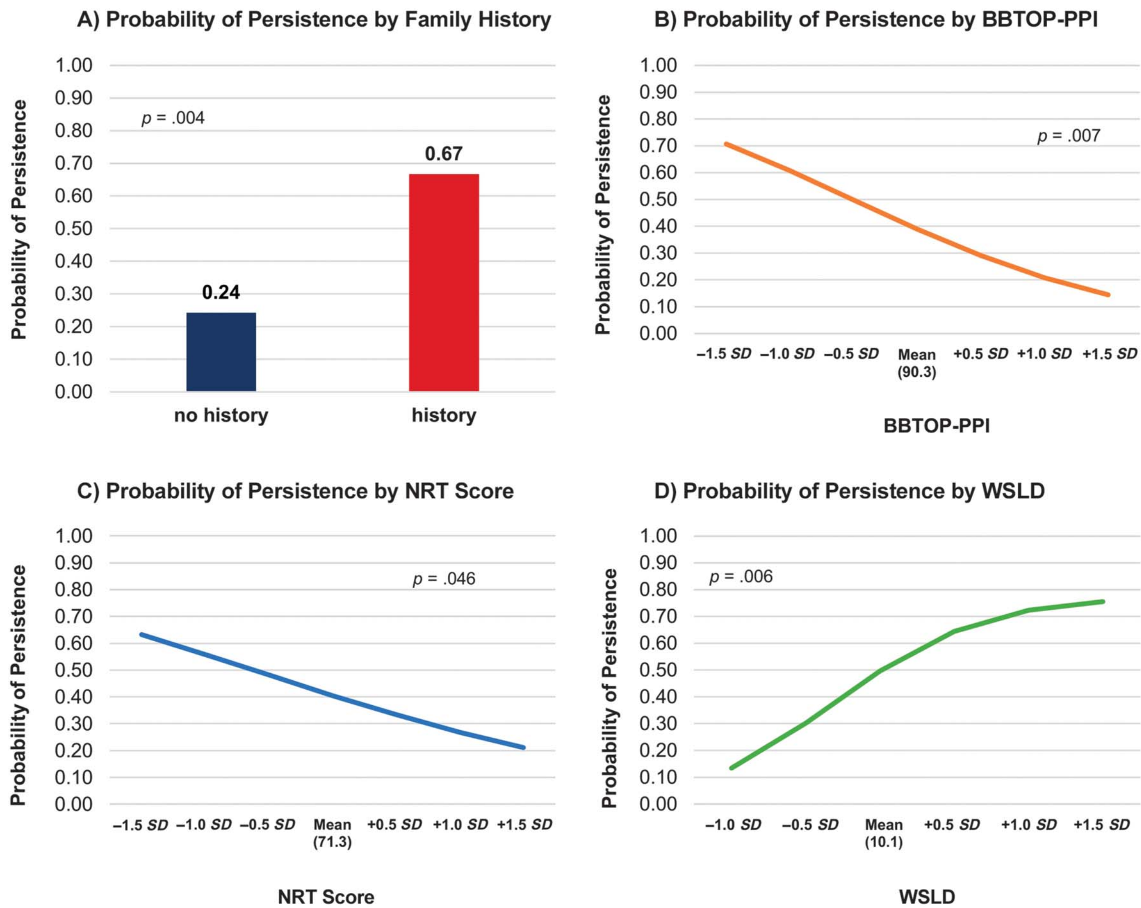
The average age at stuttering onset in months for CWS-ePer was M = 37.04, SD = 9.26, and for CWS-eRec

Table 1. Bivariate linear (and nonlinear) logistic regression results for stuttering persistence.

Variable	*OR*	95% CI		*p* value
Sex	2.02	0.54	7.61	.30
Age	1.04	0.96	1.14	.32
Age at onset	1.04	0.97	1.11	.23
Time since onset	0.99	0.93	1.05	.70
Family history recovery	3.33	0.28	39.43	.34
Family history persistence	5.20	1.42	19.04	.01
Family history	6.25	1.77	22.09	< .01
BBTOP-CI	0.94	0.90	0.99	.01
BBTOP-PPI	0.94	0.90	0.98	< .01
NRT percent	0.95	0.90	1.00	.05
WSLD	1.09	1.00	1.18	.05
Nonlinear WSLD	1.19	1.05	1.35	< .01
Nonlinear WSLD-squared	0.99	0.99	1.00	.06

Note. *OR* = odds ratio; CI = confidence interval; BBTOP-CI = Bankson–Bernthal Test of Phonology, Consonant Inventory; BBTOP-PPI = Bankson–Bernthal Test of Phonology, Phonological Process Inventory; NRT = nonword repetition test; WSLD = weighted stuttering-like disfluency.

Figure 1. Odds ratios from the bivariate regression analysis were converted into probabilities of stuttering persistence (*y*-axis) in Graphs A through D. Graph A shows positive family history (red bar) and negative family history (blue bar) as a function of the probability of stuttering persistence. Graph B shows BBTOP-PPI score by the probability of stuttering persistence. Graph C shows NRT score by the probability of persistence. Graph D shows the nonlinear effect of WSLD by the probability of persistence. NRT = nonword repetition test; WSLD = weighted stuttering-like disfluency; BBTOP-PPI = Bankson–Bernthal Test of Phonology, Phonological Process Inventory.



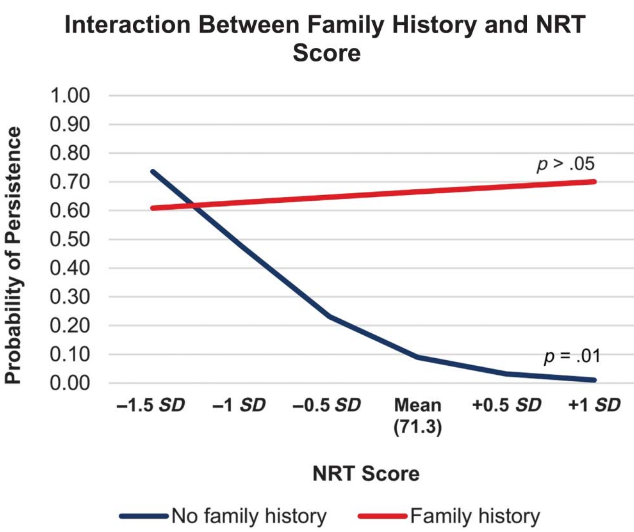
was $M = 34.16$, $SD = 7.93$. Age at stuttering onset was not associated with eventual status in the bivariate model (see Table 1). We also examined the reported length of time children had been stuttering when they entered the study. The two groups, on average, had been stuttering for comparable lengths of time, CWS-ePer $M = 18.48$, $SD = 8.68$; and CWS-eRec $M = 19.42$, $SD = 9.11$. Duration of stuttering was also not associated with eventual status in the bivariate model (see Table 1).

Although not a variable of interest in the bivariate models, we assessed whether there was a significant difference in age at the final visit of the study (when final status was determined). The average for CWS-ePer was $M = 87.33$, $SD = 13.88$; and CWS-eRec $M = 92.64$, $SD = 12.45$. The age difference between the two groups was not statistically significant, $t(52) = -1.44$, $p = .17$.

Family History

Family history of stuttering was available for 51 children (20 CWS-ePer and 31 CWS-eRec). Overall, 80.6% of CWS-eRec reported no family history of stuttering, while half as many, 40%, of CWS-ePer reported no family history. Approximately 19.4% of CWS-eRec reported a positive family history of stuttering. Of these six CWS-eRec, one out of six reported that their family member recovered, while five out of six reported their family member persisted. Sixty percent of CWS-ePer had a positive family history of stuttering, although the ratio of relatives recovering or persisting was the same as the ratio reported by CWS-eRec (i.e., two out of 12 family members recovered, and 10 out of 12 persisted). A family history of stuttering persistence and positive family history overall (regardless of whether the family member recovered or persisted) were each associated with participants' stuttering persistence. Children with a family history of persistence were 5.2 ($p = .01$) times as likely to persist versus recover, and children with a family history of stuttering (persistent or recovered) were 6.25 ($p = .004$) times as likely to persist versus recover, respectively (see Figure 1, Graph A; and Table 1). The overall family stuttering history (regardless of stuttering persistence)

Figure 2. Two-way interaction between family history and NRT score by the probability of stuttering persistence. Children with a positive family history of stuttering are represented by the red line. Children with a negative family history are represented by the blue line. NRT = nonword repetition test.



was a more powerful predictor and was therefore chosen as the stuttering history factor in subsequent models. Although we detected a large effect, the estimate is imprecise given the wide CI (see Table 1). However, the CI indicates that it is likely to be at least ~1.8 in the population and at most around 22 for positive family history.

Speech/Language Assessments

On average, CWS-ePer (M = 85, SD = 12.15) had lower BBTOP-CI scores than CWS-eRec (M = 95.16, SD = 13.73). CWS-ePer (M = 83.24, SD = 12.95) also had lower BBTOP-PPI scores compared to CWS-eRec (M = 95.16, SD = 14.22). BBTOP-CI and BBTOP-PPI scores were both associated with stuttering persistence. A one-unit increase or improvement in either score was associated with a 6% decrease in the odds of persistence versus recovery (p = .013 for BBTOP-CI, p = .007 for BBTOP-PPI; see Figure 1, Graph B; and Table 1). These two scores were highly correlated, $r(52)$ = .87, p = <.001; thus, the BBTOP-PPI measure was selected for testing in the multiple variable models because it had a more precise estimate and because of the relationship between phonological disorders and early childhood stuttering (Bloodstein & Bernstein Ratner, 2008; Wolk et al., 1993). Approximately 62% of CWS-ePer displayed delayed phonological skills at the time of testing scoring ≥ 1 SD below the normative mean (100) on this measure, while only 32% of CWS-ePer scored ≥ 1 SD below the mean on the BBTOP-PPI scale.

The average NRT score, for CWS-ePer (M = 66.95, SD = 13.47), was lower than the average NRT score for CWS-eRec (M = 74.24, SD = 10.83). Nine children scored ≥ 1 SD below the mean NRT score (M = ~71%). Six of these children were CWS-ePer, and three children were CWS-eRec. Nine children showed better performance scoring greater than 1 SD above the mean, Interestingly, three of these children were CWS-ePer and six were CWS-eRec. NRT score was associated with stuttering persistence where every one-unit increase in NRT score was associated with a 5% decline (p = .05) in the odds of persistence versus recovery (see Figure 1, Graph C; and Table 1). For comparison, our earlier study revealed that 25 age-matched CWNS achieved an average NRT score of ~77% (Spencer & Weber-Fox, 2014).

Finally, the WSLD score capturing the type and frequency of SLDs and the number of repetitions that children made during a speech sample taken when they entered the study was higher, on average, for CWS-ePer (M = 12.46, SD = 8.32) than CWS-eRec (M = 7.55, SD = 7.46). Recall that higher WSLD scores signify increased stuttering. Most children (n = 34) fell within the range of mild stuttering on the WSLD when they entered the study. Approximately 71% of children with mild stuttering were CWS-eRec. Sixteen children's WSLD scores were within the moderate range of stuttering, with 63% of these cases being CWS-ePer. WSLD scores for two children put them in the severe range of stuttering—one child was a CWS-eRec and the other a CWS-ePer. The WSLD was associated with persistence where a one-unit increase in this score was associated with a 9% increase in the likelihood of stuttering persistence (p = .046). Quadratic (nonlinear) effects were relevant for the WSLD where, after including the squared value, the linear association effect increases substantially (OR = 1.19; p = .006; see Table 1). The probability of stuttering persistence increases as WSLD increases, but this effect starts to attenuate at +0.5 SD above the mean WSLD (see Figure 1, Graph D).

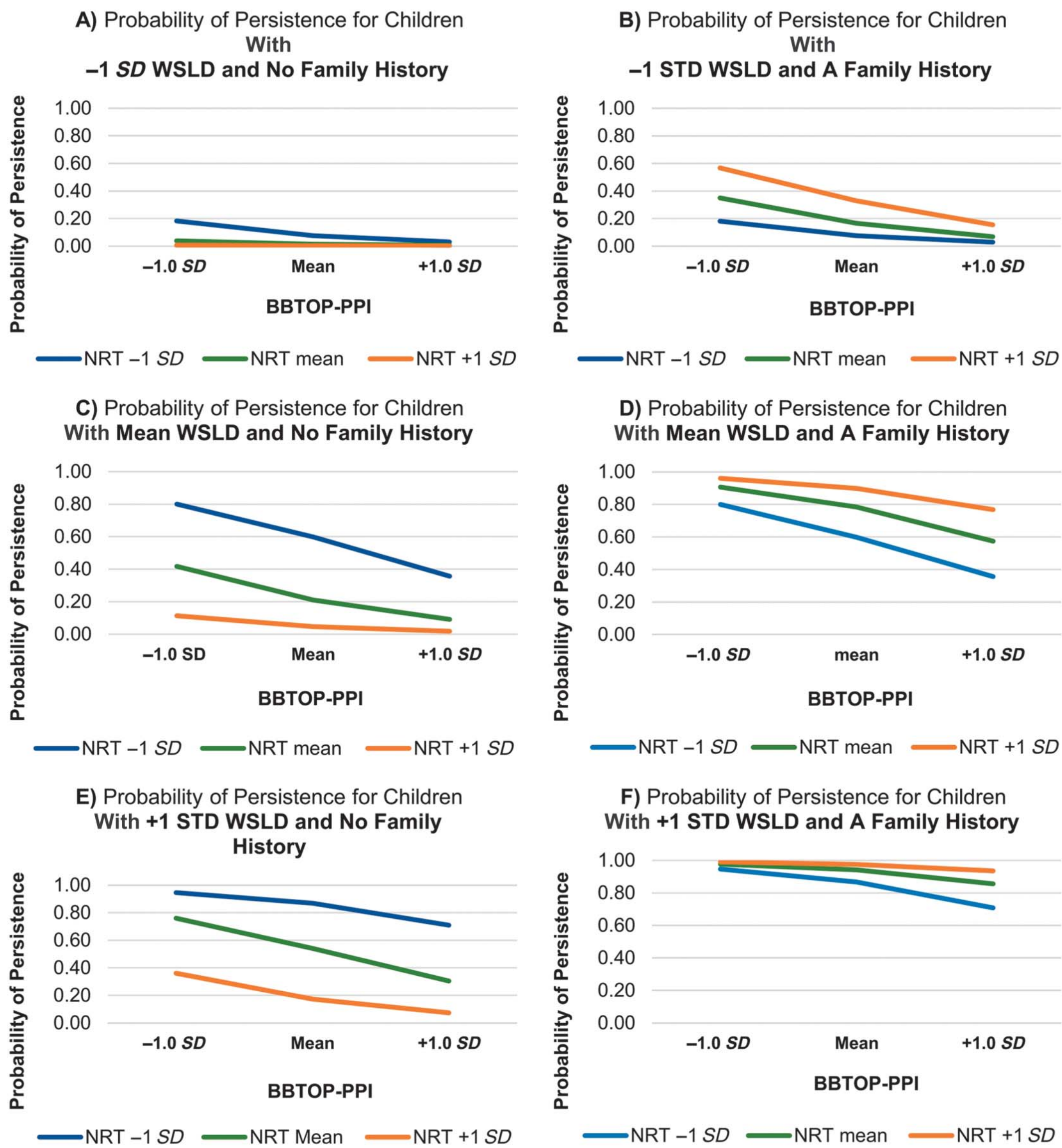
Two-Way Interactions

There was a significant interaction between overall family history and NRT score (see Table 2 and Figure 2). NRT score was associated with stuttering persistence for children without a family history but not for children with a positive family history of stuttering. For children with no family history, the OR is 0.83, indicating a 17% decline in the odds of persistence versus recovery for every 1 % increase on the NRT (see Table 2). Figure 2 reveals that the effect is strongest for NRT scores falling below the mean.

Full Model

Results from the full model with multiple predictor factors as described in Equation (1) are presented in Table 3 and illustrated in Figure 3 (nearly identical results were obtained when using the BBTOP-CI rather than the BBTOP-PPI). Most of the factors that were predictive in the sparser bivariate and interaction models remain statistically significant in the full model except for the BBTOP-PPI and the NRT percent. This was primarily due to the shared variance among these two assessments. Nevertheless, the ORs for

Figure 3. These six graphs collectively depict the full multiple variable regression model incorporating all four factors that significantly predicted stuttering persistence. The *y*-axis in each graph shows the probability of stuttering persistence, while the *x*-axis in each graph shows BBTOP-PPI score (*M* = 90.3). Graphs A, C, and E in the first column, represent no family history of stuttering, while graphs B, D, and F in the second column represent a positive family history of stuttering. The WSLD score is represented by row with the top row (Graphs A and B) representing mild stuttering (–1 *SD* below the mean), the middle row (Graphs C and D) average stuttering (*M* = 10.1), and the bottom row (Graphs E and F), more severe stuttering (+1 *SD* above the mean). NRT score is represented in each graph by the different colored lines, with orange representing better than average (+1 *SD*) NRT performance, green representing average (*M* = 71.3) NRT performance, and blue representing poorer-than-average (–1 *SD*) NRT performance. WSLD = weighted stuttering-like disfluency; NRT = nonword repetition test; BBTOP-PPI = Bankson–Bernthal Test of Phonology, Phonological Process Inventory.



these factors in the full model are similar to those found in the simpler models (see Table 1 for BBTOP-PPI and Table 2 for NRT score and family history). Predicted probabilities of stuttering persistence based on the full model are shown in Figure 3 (Graphs A through F). Graph A reveals that having a low WSLD (less frequent stuttering) and no family history of stuttering predicts a low likelihood of persistence regardless of BBTOP-PPI and NRT scores. On the other hand, Graph F shows that having higher WSLD and a positive family history predicts a high likelihood of persistence regardless of BBTOP-PPI and NRT scores. In between these extremes, the BBTOP-PPI is predictive particularly for children with average WSLD (Graphs C and D). An overall decline in the probability of persistence was noted for children with higher BBTOP-PPI scores. For children without a family history of stuttering (Graph C), the decline was most apparent for

Table 2. Interaction logistic regression results for stuttering persistence.

Variable	*OR*	95% CI		*p* value
Family history	20.13	2.88	140.69	< .01
NRT score	0.83	0.72	0.96	.01
Family history by NRT score	1.22	1.03	1.45	.03

Note. *OR* = odds ratio; CI = confidence interval; NRT = nonword repetition test.

Table 4. Best BBTOP-PPI model logistic regression results for stuttering persistence.

Variable	*OR*	95% CI		*p* value
Family history	7.36	1.49	36.27	.01
WSLD	1.26	1.07	1.48	< .01
BBTOP-PPI	0.93	0.88	0.99	.03

Note. *OR* = odds ratio; CI = confidence interval; WSLD = weighted stuttering-like disfluency; BBTOP-PPI = Bankson–Bernthal Test of Phonology, Phonological Process Inventory.

children with lower NRT accuracy scores since more accurate NRT performance was already associated with a reduced probability of persisting. The NRT score differentiated among children with no family history and less severe stuttering (Graph E) where higher NRT scores were associated with a reduced likelihood of stuttering.

Submodels

The BBTOP-PPI and NRT scores accounted for shared variance in the full multiple variable model; stated differently, they overlapped in their predictive information, which resulted in declines in the statistical significance and *OR*s for those factors in the full model. Therefore, two best-fitting submodels were estimated and tested against the full model. One submodel included the BBTOP-PPI score, family history, and WSLD (see Table 4) and the other submodel included the NRT score, family history, WSLD, and interaction between family history and NRT score (see Table 5). Unlike the full model, the BBTOP-PPI and NRT score remain statistically significant in their respective submodel. While the full model did not fit the data better than the NRT submodel, it did improve on the BBTOP submodel according to the likelihood ratio tests. Nevertheless, the full model performed better than both submodels in terms of diagnostic accuracy presented in the next section.

Model Diagnostic Accuracy

For the single variable models, submodels, and full model, we compared each participant's prediction for persistent stuttering based on probability thresholds from most to least conservative cutoffs of 0.30, 0.40, 0.50, and 0.60 with their actual persisting/recovered stuttering status at the end of the study. For each threshold cutoff, we calculated sensitivity, specificity, and positive and negative prediction values. We found that single or multiple variable models that identified children whose risk for persistence was 40% or higher as candidates for immediate intervention resulted in better diagnostic validity than other cutoff values for risk of persistence. Thus, we used this 0.40 to report the four accuracy measures for each model in Figure 4. For reference, model results for other cutoff scores are provided in Appendix B. From Figure 4, we see that the multiple variable models resulted in higher accuracy compared to the single variable models revealed by their sensitivity, specificity, and positive and negative prediction values, with the full model outperforming the multiple variable submodels. Figure 5 provides the error rates, or the total number of false positives (children incorrectly identified as persisting) and false negatives (children incorrectly identified as recovered) out of the total number of participants, for each model at cutoffs of 0.30, 0.40, 0.50, and 0.60. The multiple variable models resulted in lower error rates compared to the single variable models with the full model outperforming the other models with the lowest error rates. Specifically, thresholds of 0.40 and 0.50 resulted in the lowest error rates for the full model (see Figure 5).

Table 3. Full multiple variable logistic regression results for stuttering persistence.

Variable	*OR*	95% CI		*p* value
Family history	13.60	1.55	119.32	.02
BBTOP-PPI	0.94	0.86	1.02	.13
NRT score	0.87	0.74	1.01	.07
Family history by NRT score	1.24	1.02	1.51	.03
WSLD	1.31	1.04	1.65	.02
WSLD-squared	0.99	0.98	1.00	.09

Note. *OR* = odds ratio; CI = confidence interval; BBTOP-PPI = Bankson–Bernthal Test of Phonology, Phonological Process Inventory; NRT = nonword repetition test; WSLD = weighted stuttering-like disfluency.

Discussion

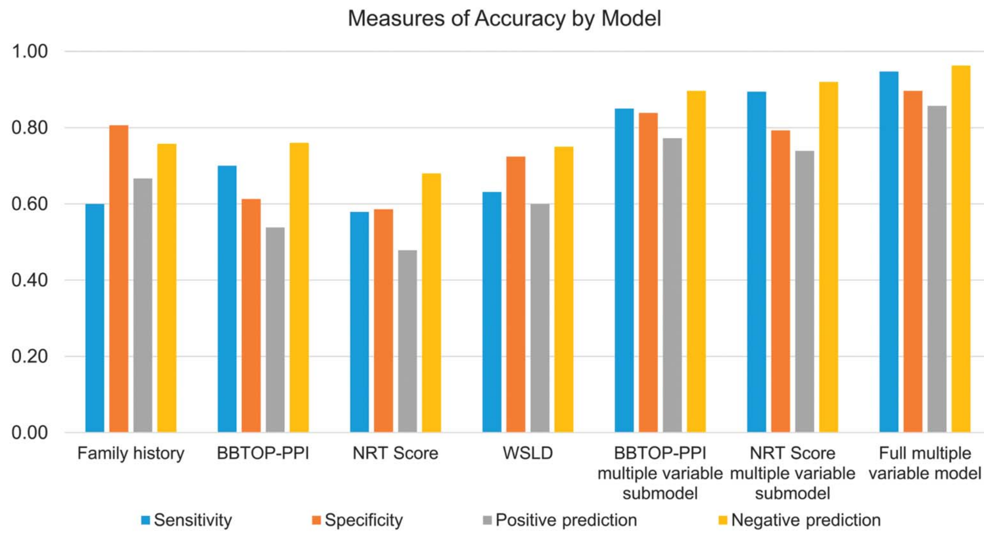
The importance of assessing prognostic factors that place a child at greatest risk for stuttering persistence has long been recognized (Yairi & Ambrose, 2005). However, there

Table 5. Best NRT score model logistic regression results for stuttering persistence

Variable	*OR*	95% CI		*p* value
Family history	21.45	2.43	189.34	< .01
WSLD	1.20	1.01	1.43	.04
NRT score	0.82	0.70	0.97	.02
Family history by NRT score	1.25	1.03	1.52	.03

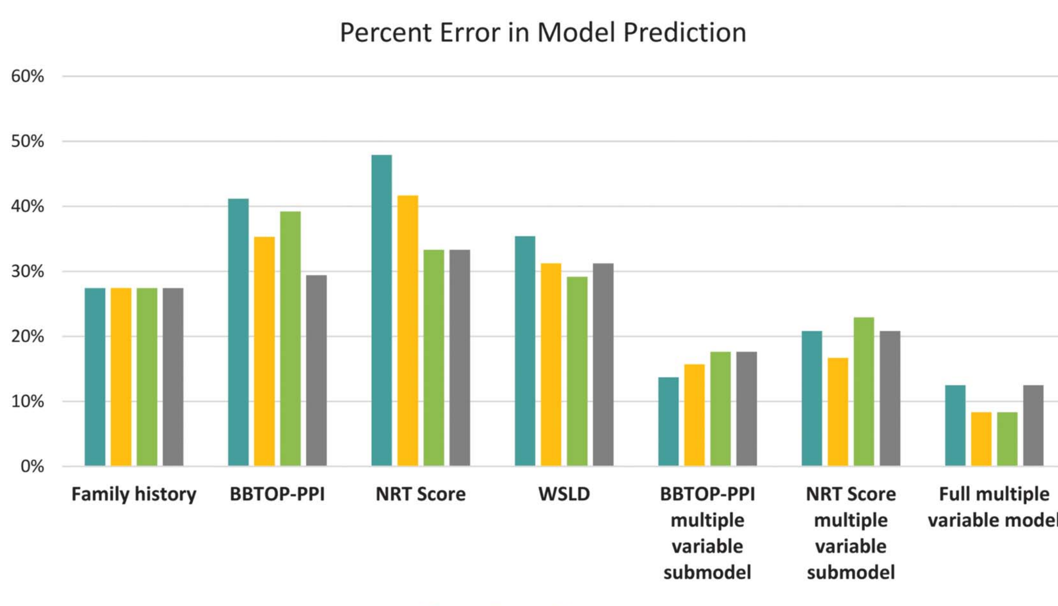
Note. *OR* = odds ratio; CI = confidence interval; WSLD = weighted stuttering-like disfluency; NRT = nonword repetition test.

Figure 4. Accuracy (*y*-axis) for the single and multiple variable regression models (*x*-axis). Sensitivity (true positive/false negative) is the blue bar, specificity (false positive/true negative) is the orange bar, positive prediction (true positive/false positive) is the gray bar, and negative prediction (false negative/true negative) is the yellow bar. NRT = nonword repetition test; WSLD = weighted stuttering-like disfluency; BBTOP-PPI = Bankson–Bernthal Test of Phonology, Phonological Process Inventory.



have been few studies examining multiple risk factors for stuttering persistence in the same group of children (Ambrose et al., 2015). This is the first study, to our knowledge, that assesses the complex relationships among multiple risk factors for stuttering persistence to establish whether multiple variable models predict higher risk for eventual stuttering persistence. We also provide measures of diagnostic accuracy for each model at different cutoff scores to guide SLPs in their clinical decision making. Results from the bivariate logistic regression analysis confirmed that a positive family history of stuttering, lower speech sound production accuracy measured by the BBTOP-PPI (and CI) scales, the type

Figure 5. Error rate (*y*-axis) or the total number of errors (false positives plus false negatives) out of the total number of subjects for the single and multiple variable regression models (*x*-axis). Colored bars represent cutoff levels from more to less conservative: .30 (turquoise), .40 (yellow), .50 (green), and .60 (gray). NRT = nonword repetition test; WSLD = weighted stuttering-like disfluency; BBTOP-PPI = Bankson–Bernthal Test of Phonology, Phonological Process Inventory.



and frequency of SLDs as measured by the WSLD, and poorer nonword repetition performance each predicted stuttering persistence and recovery in our sample of preschool CWS. However, the multiple variable regression model that included all these factors resulted in the best fitting model with the highest sensitivity, specificity, positive and negative predictive values, and subsequent lowest error rate.

Risk Factors for Persistent Stuttering

Epidemiological Factors

We conducted binary logistic regression to determine whether there were statistically significant relationships between epidemiological, stuttering frequency, or linguistic risk factors and stuttering outcome—persistence or recovery. Corroborating earlier studies in preschoolers, we found that CWS-ePer were more likely to have a positive family history of stuttering (Yairi & Ambrose, 1992a; Yairi et al., 1996). Children with a family history had a 67% chance of persisting, while children with no family history had only a 24% chance of persisting (see Figure 1, Graph A). Unlike Yairi et al. (1996) who found that children were more likely to recover or persist based on their relative's status (i.e., whether they recovered or persisted themselves), we found that a positive family history, regardless of the family member's outcome, was the strongest predictor of stuttering outcome confirmed by the *OR* in Table 1.

We did not find a relationship between sex and stuttering outcome, which seems at odds with the established relationship between sex and persistence. However, given that we had far fewer females in the study—and only four in the persistent group—we had insufficient power to detect potential differences. As Yairi and Ambrose (1999) noted, the fact that far fewer girls persist "makes it difficult to test the significance of [this] infrequency" (p. 1104). The recent meta-analysis of factors related to persistence and recovery revealed that boys were 1.48 times as likely to persist in stuttering (Singer et al., 2020).

Age at stuttering onset did not predict persistence. On average, CWS-ePer began stuttering approximately 3 months later than CWS-eRec. This is comparable to the 3.5 month later age at onset for CWS-ePer reported by Yairi and Ambrose (2005). There was considerable overlap between the groups, however, and the difference in age at onset between CWS-eRec and CWS-ePer in our study (and in Yairi and Ambrose's study) did not reach statistical significance. Singer et al. (2020) integrated data from six studies found that CWS with later age at onsets were more likely to persist; however, they did not clarify what necessarily constitutes a later age at stuttering onset.

We also did not find that the duration a child had been stuttering to be predictive of eventual outcome. Yairi and Ambrose (2005) found that a child's chance of recovery decreases after approximately 1 year of stuttering and documented a 75%–80% recovery rate within 6–15 months of stuttering onset. The children in this study had been stuttering for well over a year (CWS-ePer M = 18.5 months; CWS-eRec M = 19.4 months) when they entered the study, and as expected, we noted a lower, 59.6%, probability of recovery. Thus, the duration of stuttering in children who are, on average, 54 months may not be as useful a predictor as it is for CWS at other ages.

Stuttering Frequency and Linguistic Factors

We anticipated that BBTOP, NRT, and WSLD scores would significantly predict stuttering persistence because data for this study overlapped with data in our earlier reports (Spencer & Weber-Fox, 2014; Walsh et al., 2018, 2020). The co-occurrence of speech sound disorders and stuttering in preschoolers has long been recognized, although estimates of CWS with these concomitant deficits vary considerably from as low as 7% in a community cohort sample (Unicomb et al., 2020) to as high as 46% in a survey of practicing clinicians (Blood et al., 2003). The Illinois studies documented speech and language development in children near the onset of stuttering through the preschool years (Yairi & Ambrose, 2005). They found that CWS-ePer lagged CWS-eRec in phonological development, although children persisting in stuttering eventually caught up with their peers who recovered and acquired phonological skills in the typical developmental sequence (Paden et al., 1999). The recent meta-analysis (that included data from the Purdue and Illinois studies) based on BBTOP, Goldman-Fristoe Test of Articulation (Goldman & Fristoe, 2000), and the Assessment of Phonological Processes–Revised (Hodson, 1986) confirmed reduced speech sound production accuracy as a risk factor for stuttering persistence (Singer et al., 2020). Children who scored 1 *SD* below the group average, ~90, on the BBTOP-PPI in our study had an approximately 57% risk of persistence, while scoring 1 *SD* above the group average yielded a much lower, ~22%, risk for persistence (see Figure 1, Graph B).

NRTs have been used to assess phonological abilities in CWS, specifically, the perception, encoding, retention, and execution of novel phonemic sequences. Studies reveal less accurate performance achieved by CWS compared to CWNS, particularly for longer nonwords (Anderson et al., 2006; Anderson & Wagovich, 2010; Hakim & Ratner, 2004; Pelczarski & Yaruss, 2016). An earlier study from our lab found that children who would eventually persist performed with poorer accuracy on the NRT compared to children who would eventually recover from stuttering (Spencer & Weber-Fox, 2014). This study based on these same data revealed that the average NRT score was approximately 71%, and children scoring below this average had a greater risk of persisting (see Figure 1, Graph C). However, this finding is complicated by the significant interaction we detected between family history and NRT score (see Figure 2). From this figure, we see that NRT scores for children with a positive family history, even at the extremes in the range of scores, do not offer a significant increase (increased precision) in prediction of persistence. For children without a family history of stuttering, however, there is a clear relationship between NRT performance and the risk of persistence. Children with no family history scoring 1 *SD* or more below the average NRT score had a 50%–73% probability of stuttering persistence, while children performing at the mean or better had lower probabilities of persisting. Thus, for children without a positive family history of stuttering, the NRT task, which can be administered and scored in a reasonable amount of time, may be a useful additional measure to help predict a child's overall risk for stuttering persistence.

Finally, the WSLD score considers the type, frequency, and number of repetitions of SLDs, with higher scores indicating more frequent/severe stuttering. As the curve in Figure 1, Graph D suggests, there was a nonlinear relationship between the WSLD and persistence. Higher WSLD scores were associated with a higher probability of persistence. On average, children in the study scored a 10.1. Children earning a score 1 *SD* above this average placed them at an ~80% chance of persisting, while a score 1 *SD* below the average placed them at only a 10% chance of persisting. In our recent study, Walsh et al. (2020), we discussed that although CWS-ePer, on average, did not show the same rate of decline in SLDs as CWS-eRec (Yairi & Ambrose, 1992b; Yairi et al., 1996), the difference between the two groups did not reach significance in children under age 4 years (Throneburg & Yairi, 2001; Yairi & Ambrose, 1992b). However, our study with slightly older, on average, preschoolers revealed significant differences in stuttering frequency, indexed by the WSLD, between children who would eventually recover or persist (Walsh et al., 2020).

Multiple Variable Models

We hypothesized that models incorporating a combination of risk factors would result in more accurate predictions of stuttering persistence compared to sparser models. Although intuitive, this is the first study to lend empirical support to this assumption by examining multiple predictive risk factors within the same child who stutters. We discovered that the full model that considers family history, BBTOP-PPI score, NRT accuracy, and WSLD scores resulted in the most accurate predictions indexed by higher sensitivity, specificity, and positive and negative prediction values. The full model resulted in the lowest error rates that combined false positives and false negatives predicted by each model.

Each CWS has a unique profile of risk factors that contribute to the probability of that child persisting or recovering from early childhood stuttering; thus, it is important to assess multiple factors for a child who is stuttering. A general guideline is for clinicians to prioritize treatment services for those children manifesting a greater number of risk factors for stuttering persistence (Conture, 2001; Yairi & Ambrose, 2005; Yaruss & Reardon-Reeves, 2017). Our full model provides, for the first time, empirical evidence that multiple risk factors should indeed be considered in combination where different factors combine in unique ways to predict stuttering. Some combinations of factors dominate the prediction. For example, at the extremes, the full model revealed that children with mild stuttering and no family history had an extremely low probability of persisting, whereas children with more frequent stuttering and a positive family history were most likely to persist in stuttering. In these cases, BBTOP and NRT performance did not add substantially to the probabilities of persistence as these probabilities were already high or low, respectively. In between these extremes, the BBTOP and NRT offer prognostic value. Average and above average BBTOP-PPI scores predicted a lower risk of persisting in children with mild-to-moderate levels of stuttering, while NRT performance may be a useful predictor for those children without a family history of stuttering.

Clinical Implications and Key Takeaways

To enhance the translational value of our study, we computed accuracy measures for the predictive statistical models to determine their precision in diagnosing stuttering persistence and recovery for each child. When assessing a 3- to 5-year-old CWS like the children in our sample, SLPs can refer to the graphs in Figure 3 to approximate that child's risk for persistence by determining which graph best captures their profile of risk factors. For example, if a CWS does not have a family history of stuttering, an SLP can use the three graphs in the first column (A, C, and E) to determine where that child may fall with respect to the other factors. If the child had a WSLD score around the average for this sample (10.1), then the NRT could be administered as an additional assessment tool with poorer performance indicating greater risk (Graph C). On the other hand, if a child has a positive family history, the graphs in the second column (B, D, and F) could be used to determine where that child falls with respect to the other factors.

We provided four diagnostic accuracy indices for each model (see Figure 4) along with error rates for different cutoff scores (see Figure 5). This step was undertaken to demonstrate how these methods might be used with future samples of children to determine optimal cutoff scores. We would argue for adopting more stringent cutoffs, thus prioritizing sensitivity over specificity when making consequential decisions about whether to recommend immediate treatment for a child who is stuttering. In Walsh et al. (2020), we noted that, "In the case of early intervention, failing to identify a [true positive] could have profound lifelong ramifications…[while] recommending treatment for a child who would have recovered without treatment, albeit costly and an expenditure of scarce resources, may be a more acceptable trade-off with less impactful consequences." (pp. 2562–2563).

A family history of stuttering—regardless of whether a family member persisted or recovered—is a powerful predictor, revealed by the higher *OR*s in Tables 1–3, of persistence and indicator for immediate intervention. Thus, it is well-worth consulting with caregivers to determine whether parents, siblings, grandparents, aunts, uncles, great-grandparents, and first cousins currently or ever stuttered. The frequency and severity of SLDs, indexed by the WSLD, is another strong prognostic indicator of persistence. We acknowledge that recording a speech sample and calculating the WSLD is time consuming; however, Singer et al. (2020) found that clinicians and parents' scaled stuttering severity ratings did not differentiate between children who would eventually recover or persist. Walsh et al. (2018) found a moderate correlation between the WSLD and a standardized assessment, the Test of Childhood Stuttering (TOCS; Gillam et al., 2009). Although we do not have TOCS scores for this cohort of children, future studies could determine the accuracy with which TOCS scores predict persistence as a potential alternative to the WSLD.

If a concomitant speech sound disorder is suspected, then administering an assessment of phonology/articulation is a key component of a comprehensive assessment that may contribute additional predictive value. Although we showed that better performance on the BBTOP (i.e., 1 *SD* above the mean of 91) was a positive sign for recovery, we would not recommend administering a phonological/articulation assessment if a child is showing typical development given the significant constraints on clinicians' time to conduct evaluations. Administering and scoring these assessments is time intensive, and SLPs could ultimately estimate a child's performance in the graphs to be at/above the mean if there are no concerns about concomitant speech sound disorders. On the other hand, the Dollaghan NRT can be administered and scored in approximately 10 min and may serve as an additional prognostic indicator of persistence. The NRT and BBTOP are likely to be assessing common underlying phonological processes as these factors' *p* values were no longer significant when entered together in the full model due to shared variance (although the *OR*s for these factors were comparable to the *OR*s in the single variable models). However, these measures may access distinct phonological processes as well. Spencer and Weber-Fox (2014) reported that NRT and BBTOP-CI scores were moderately correlated and that the NRT predicted persistence and recovery even when CWS with concomitant language and/or speech sound disorders were omitted from their analysis. Moreover, the full model, which included both measures, outperformed all the other models in terms of predictive accuracy.

Our predictive models give clinicians probabilities of stuttering persistence and recovery in preschool children based on multiple risk factors. A question to consider is how SLPs might use this information to counsel parents to help them make consequential decisions regarding therapy. First, when discussing a child's risk profile, it is critical to emphasize to parents that risk factors give us likelihoods of stuttering persistence and recovery. They are our best estimates of future outcomes, yet they are not definitive indicators of whether their child's stuttering will indeed persist or resolve. Each child is unique and may show an unexpected pattern—a child at high risk could recover, while a child at low risk could persist. If a child is at high risk for persistence, SLPs should assure parents that it does not definitively indicate that their child's stuttering will persist, but in the meantime, they will work with the family and child toward the goal of helping that child become a successful communicator. If parents decide to wait to initiate therapy, for example, if their child presents a low risk for stuttering persistence, it is important that the parent and SLP carefully continue to monitor the child. Parents should be encouraged to follow up with the SLP at regular intervals and to contact the SLP if they notice changes in their child's speech. Finally, regardless of their risk profile, if a child (or their parent) is expressing concern, anxiety, or negative feelings and attitudes toward their communication abilities, that child (and family) would clearly benefit from intervention addressing the impactful issues that often accompany the stuttering condition.

Limitations and Considerations

This study has several limitations and other considerations warranting thoughtful discussion. First, our models were based on data from a moderate sample size of 52 preschoolers who stutter. These models will clearly be enhanced by the inclusion of more children's data. We aim to broaden our approach to include a larger sample of CWS at different ages and incorporate additional risk measures (Singer et al., 2020). A future goal of this work is to expand this database and create an online calculator in which clinicians can enter a child's age, relevant epidemiological factors, and other assessment scores collected during an evaluation to determine an individual child's collective risk for stuttering persistence. Nevertheless, our sample of 52 CWS demonstrated a range in performance on the clinical measures likely seen in the population at large. CWS showed delayed or advanced speech sound production abilities and nonword production performance. Children had positive or negative family histories of stuttering and displayed a range of stuttering frequency/WSLD scores from mild to moderately severe. Second, some children were not followed as long as others, so these children's stuttering status may have been in a state of transition. This issue could have affected group classification, which was based on the data that we had for these children at the time. Third, it is important to limit inference to CWS within the age boundaries (i.e., 3–5 years) of the population represented in this sample (see details in Appendix A). This study is a critical first step establishing our methods to reveal how multiple factors can be used to predict collective risk of stuttering persistency. We would encourage clinicians to use the predictive values in the tables and Figure 3 to guide their assessment of a child who is stuttering, provided that the child is between the ages of 3–5 years with typical nonverbal intelligence and no indicators of autism spectrum disorders (per our inclusionary criteria). It would, however, not be prudent at this stage to generalize findings to CWS at other ages, as the risk factors found to be predictive in our sample of children may not necessarily be predictive for children at other ages. It may be the case that estimators of risk change with age. For example, stuttering severity is not predictive of eventual persistence or recovery in younger preschoolers aged 2–4 years (Throneburg & Yairi, 2001; Yairi & Ambrose, 1992b). Exploring other risk factors and incorporating data from CWS at younger/older ages is an important next step for this research to establish models that integrate risk factors for stuttering persistence for children across the developmental window of early childhood.

Conclusions

For the first time, we show how factors collectively predict the probability of persistence in preschoolers who are stuttering. Consistent with the MDP theory of stuttering (Smith & Weber, 2017), we found that using multiple risk factors to assess children yielded more accurate predictions of stuttering persistence and recovery compared to factors in isolation. Although we focused on risk factors accessible to clinicians, research from our lab and others has revealed critical neurophysiological risk factors for stuttering persistence such as trajectories of neuroanatomical development (Chow & Chang, 2017; Garnett et al., 2018), patterns of neural activation during speech (Hosseini et al., 2018), articulatory kinematics (Usler et al., 2017), and sympathetic nervous system arousal (Zengin-Bolatkale et al., 2018) associated with stuttering persistence and recovery. A better understanding of the diverse neurophysiological, epidemiological, and clinical factors that underlie stuttering persistence and recovery will yield insight into the underpinnings of chronic stuttering and will help identify etiological targets for novel interventions.

Acknowledgments

This research was supported by the National Institute on Deafness and Other Communication Disorders (R01DC00559 awarded to Anne Smith and Christine Weber and R01DC018000 awarded to Bridget Walsh). We wish to acknowledge Anne Smith, Distinguished Professor Emerita from Purdue University, whose research in stuttering spanned 3 decades—mentoring future generations of researchers in the field and laying the foundation for multifactorial approaches to understand the development of stuttering in young children. We thank Barb Brown and Janna Berlin for their many years of dedication to the Purdue Stuttering Project. We also thank Patricia Zebrowski and her colleagues at The University of Iowa for collaborating and helping us acquire a portion of this data. Lastly, we are grateful to all the children and their families whose participation in our research made this study possible.

References

Ambrose, N. G., & Yairi, E. (1999). Normative disfluency data for early childhood stuttering. *Journal of Speech, Language, and Hearing Research, 42*(4), 895–909. https://doi.org/10.1044/jslhr.4204.895

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

Anderson, J. D., & Wagovich, S. A. (2010). Relationships among linguistic processing speed, phonological working memory, and attention in children who stutter. *Journal of Fluency Disorders, 35*(3), 216–234. https://doi.org/10.1016/j.jfludis.2010.04.003

Anderson, J. D., Wagovich, S. A., & Hall, N. E. (2006). Nonword repetition skills in young children who do and do not stutter. *Journal of Fluency Disorders, 31*(3), 177–199. https://doi.org/10.1016/j.jfludis.2006.05.001

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

Bauman, J., Hall, N. E., Wagovich, S. A., Weber-Fox, C. M., & Bernstein Ratner, N. (2012). Past tense marking in the spontaneous speech of preschool children who do and do not stutter. *Journal of Fluency Disorders, 37*(4), 314–324. https://doi.org/10.1016/j.jfludis.2012.04.003

Blood, G. W., Ridenour, V. J., Qualls, C. D., & Hammer, C. S. (2003). Co-occurring disorders in children who stutter. *Journal of Communication Disorders, 36*(6), 427–448. https://doi.org/10.1016/S0021-9924(03)00023-6

Bloodstein, O., & Bernstein Ratner, N. (2008). *A handbook on stuttering* (6th ed.). Delmar Learning.

Blumgart, E., Tran, Y., & Craig, A. (2010). Social anxiety disorder in adults who stutter. *Depression and Anxiety, 27*(7), 687–692. https://doi.org/10.1002/da.20657

Burgemeister, B., Blum, L., & Lorge, I. (1972). *The Columbia Mental Maturity Scale*. Harcourt Brace Jovanovich.

Carrow-Woolfolk, E. (1999). *Test for Auditory Comprehension of Language–3rd Edition*. AGS.

Chang, S.-E., & Zhu, D. C. (2013). Neural network connectivity differences in children who stutter. *Brain, 136*(12), 3709–3726. https://doi.org/10.1093/brain/awt275

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

Constantino, C. D., Leslie, P., Quesal, R. W., & Yaruss, J. S. (2016). A preliminary investigation of daily variability of stuttering in adults. *Journal of Communication Disorders, 60,* 39–50. https://doi.org/10.1016/j.jcomdis.2016.02.001

Conture, E. G. (2001). *Stuttering: Its nature, diagnosis and treatment*. Pearson College Division.

Craig, A., Blumgart, E., & Tran, Y. (2009). The impact of stuttering on the quality of life in adults who stutter. *Journal of Fluency Disorders, 34*(2), 61–71. https://doi.org/10.1016/j.jfludis.2009.05.002

Dawson, J., Stout, C., & Eyer, J. (2003). *Structured Photographic Expressive Language Test–Third Edition (SPELT-3)*. Janelle Publications.

De Nil, L. F., & Brutten, G. J. (1990). Speech-associated attitudes. *Journal of Fluency Disorders, 15*(2), 127–134. https://doi.org/10.1016/0094-730X(90)90039-U

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*(5), 1136–1146. https://doi.org/10.1044/jslhr.4105.1136

Dworzynski, K., Remington, A., Rijsdijk, F., Howell, P., & Plomin, R. (2007). Genetic etiology in cases of recovered and persistent stuttering in an unselected, longitudinal sample of young twins. *American Journal of Speech-Language Pathology, 16*(2), 169–178. https://doi.org/10.1044/1058-0360(2007/021)

Gardner, M. (1985). *Test of Auditory-Perceptual Skills*. Children's Hospital of San Francisco.

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., Logan, K., & Pearson, N. (2009). *Test of Childhood Stuttering*. Pro-Ed.

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

Hakim, H. B., & Ratner, N. B. (2004). Nonword repetition abilities of children who stutter: An exploratory study. *Journal of Fluency Disorders, 29*(3), 179–199. https://doi.org/10.1016/j.jfludis.2004.06.001
Hodson, B. W. (1986). *The Assessment of Phonological Processes–Revised*. Interstate.
Hosseini, R., Walsh, B., Tian, F., & Wang, S. (2018). An fNIRS-based feature learning and classification framework to distinguish hemodynamic patterns in children who stutter. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26*(6), 1254–1263. https://doi.org/10.1109/TNSRE.2018.2829083
Janssen, P., Kloth, S., Kraaimaat, F., & Brutten, G. J. (1996). Genetic factors in stuttering: A replication of Ambrose, Yairi, and Cox's (1993) study with adult probands. *Journal of Fluency Disorders, 21*(2), 105–108. https://doi.org/10.1016/0094-730X(95)00060-K
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
Kreidler, K., Hampton Wray, A., Usler, E., & Weber, C. (2017). Neural indices of semantic processing in early childhood distinguish eventual stuttering persistence and recovery. *Journal of Speech, Language, and Hearing Research, 60*(11), 3118–3134. https://doi.org/10.1044/2017_JSLHR-S-17-0081
Leech, K. A., Bernstein Ratner, N., Brown, B., & Weber, C. M. (2017). Preliminary evidence that growth in productive language differentiates childhood stuttering persistence and recovery. *Journal of Speech, Language, and Hearing Research, 60*(11), 3097–3109. https://doi.org/10.1044/2017_JSLHR-S-16-0371
MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Erlbaum.
Miller, J., & Iglesias, A. (2006). *Systematic analysis of language transcripts (SALT)* (Version 9) [Computer software]. University of Wisconsin–Madison.
Mohan, R., & Weber, C. (2015). Neural systems mediating processing of sound units of language distinguish recovery versus persistence in stuttering. *Journal of Neurodevelopmental Disorders, 7*(1), Article 28. https://doi.org/10.1186/s11689-015-9124-7
Munson, B., Edwards, J., & Beckman, M. E. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech, Language, and Hearing Research, 48*(1), 61–78. https://doi.org/10.1044/1092-4388(2005/006)
Ntourou, K., Conture, E. G., & Lipsey, M. W. (2011). Language abilities of children who stutter: A meta-analytical review. *American Journal of Speech-Language Pathology, 20*(3), 163–179. https://doi.org/10.1044/1058-0360(2011/09-0102)
O'Brian, S., Jones, M., Packman, A., Menzies, R., & Onslow, M. (2011). Stuttering severity and educational attainment. *Journal of Fluency Disorders, 36*(2), 86–92. https://doi.org/10.1016/j.jfludis.2011.02.006
Paden, E. P., Yairi, E., & Ambrose, N. G. (1999). Early childhood stuttering II. *Journal of Speech, Language, and Hearing Research, 42*(5), 1113–1124. https://doi.org/10.1044/jslhr.4205.1113
Pelczarski, K. M., & Yaruss, J. S. (2016). Phonological memory in young children who stutter. *Journal of Communication Disorders, 62,* 54–66. https://doi.org/10.1016/j.jcomdis.2016.05.006
Ratner, N. B., Rooney, B., & MacWhinney, B. (1996). Analysis of stuttering using CHILDES and CLAN. *Clinical Linguistics & Phonetics, 10*(3), 169–187. https://doi.org/10.3109/02699209608985170
Reilly, S., Onslow, M., Packman, A., Cini, E., Conway, L., Ukoumunne, O. C., Bavin, E. L., Prior, M., Eadie, P., Block, S., & Wake, M. (2013). Natural history of stuttering to 4 years of age: A prospective community-based study. *Pediatrics, 132*(3), 460–467. https://doi.org/10.1542/peds.2012-3067
Reilly, S., Onslow, M., Packman, A., Wake, M., Bavin, E. L., Prior, M., Eadie, P., Cini, E., Bolzonello, C., & Ukoumunne, O. C. (2009). Predicting stuttering onset by the age of 3 years: A prospective, community cohort study. *Pediatrics, 123*(1), 270–277. https://doi.org/10.1542/peds.2007-3219
Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *Childhood Autism Rating Scale* (CARS). Western Psychological Services.
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., 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
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
Throneburg, R. N., & Yairi, E. (2001). Durational, proportionate, and absolute frequency characteristics of disfluencies. *Journal of Speech, Language, and Hearing Research, 44*(1), 38–51. https://doi.org/10.1044/1092-4388(2001/004)
Tran, Y., Blumgart, E., & Craig, A. (2011). Subjective distress associated with chronic stuttering. *Journal of Fluency Disorders, 36*(1), 17–26. https://doi.org/10.1016/j.jfludis.2010.12.003
Trevethan, R. (2017). Sensitivity, specificity, and predictive values: Foundations, liabilities, and pitfalls in research and practice. *Frontiers in Public Health, 5,* 307. https://doi.org/10.3389/fpubh.2017.00307
Unicomb, R., Kefalianos, E., Reilly, S., Cook, F., & Morgan, A. (2020). Prevalence and features of comorbid stuttering and speech sound disorder at age 4 years. *Journal of Communication Disorders, 84,* Article 105976. https://doi.org/10.1016/j.jcomdis.2020.105976
Usler, E., Smith, A., & Weber, C. (2017). A lag in speech motor coordination during sentence production is associated with stuttering persistence in young children. *Journal of Speech, Language, and Hearing Research, 60*(1), 51–61. https://doi.org/10.1044/2016_JSLHR-S-15-0367
Usler, E., & Weber-Fox, C. (2015). Neurodevelopment for syntactic processing distinguishes childhood stuttering recovery versus persistence. *Journal of Neurodevelopmental Disorders, 7*(1), Article 4. https://doi.org/10.1186/1866-1955-7-4
Vanryckeghem, M., Brutten, G. J., & Hernandez, L. M. (2005). A comparative investigation of the speech-associated attitude of preschool and kindergarten children who do and do not stutter. *Journal of Fluency Disorders, 30*(4), 307–318. https://doi.org/10.1016/j.jfludis.2005.09.003
Viswanath, N., Lee, H. S., & Chakraborty, R. (2004). Evidence for a major gene influence on persistent developmental stuttering. *Human Biology, 76*(3), 401–412. https://doi.org/10.1353/hub.2004.0050
Wagovich, S. A., Hall, N. E., & Clifford, B. A. (2009). Speech disruptions in relation to language growth in children who stutter: An exploratory study. *Journal of Fluency Disorders, 34*(4), 242–256. https://doi.org/10.1016/j.jfludis.2009.09.004

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., Tian, F., Tourville, J. A., Yücel, M. A., Kuczek, T., & Bostian, A. J. (2017). Hemodynamics of speech production: An fNIRS investigation of children who stutter. *Scientific Reports, 7*(1), 1–13. https://doi.org/10.1038/s41598-017-04357-6

Walsh, B., Usler, E., Bostian, A., Mohan, R., Gerwin, K. L., Brown, B., Weber, C., & Smith, A. (2018). What are predictors for persistence in childhood stuttering? *Seminars in Speech and Language, 39*(4), 299–312. https://doi.org/10.1055/s-0038-1667159

Wolk, L., Edwards, M. L., & Conture, E. G. (1993). Coexistence of stuttering and disordered phonology in young children. *Journal of Speech and Hearing Research, 36*(5), 906–917. https://doi.org/10.1044/jshr.3605.906

Yairi, E., & Ambrose, N. (1992a). Onset of stuttering in preschool children. *Journal of Speech and Hearing Research, 35*(4), 782–788. https://doi.org/10.1044/jshr.3504.782

Yairi, E., & Ambrose, N. G. (1992b). A longitudinal study of stuttering in children. *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. G., & Niermann, R. (1993). The early months of stuttering. *Journal of Speech and Hearing Research, 36*(3), 521–528. https://doi.org/10.1044/jshr.3603.521

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. (1997). Clinical implications of situational variability in preschool children who stutter. *Journal of Fluency Disorders, 22*(3), 187–203. https://doi.org/10.1016/S0094-730X(97)00009-0

Yaruss, J. S. (2010). Assessing quality of life in stuttering treatment outcomes research. *Journal of Fluency Disorders, 35*(3), 190–202. https://doi.org/10.1016/j.jfludis.2010.05.010

Yaruss, J. S., LaSalle, L. R., & Conture, E. G. (1998). Evaluating stuttering in young children. *American Journal of Speech-Language Pathology, 7*(4), 62–76. https://doi.org/10.1044/1058-0360.0704.62

Yaruss, J. S., & Reardon-Reeves, N. (2017). *Early childhood stuttering therapy: A practical guide*. Stuttering Therapy Resources.

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

Participant Characteristics

Subject	Status	Sex	Age (at Visit 1)	Age of onset	Duration of stuttering (at Visit 1)	Family history	Age at last visit/ last study visit
Participant 1	Per	M	48	30	18	Y	73/Yr 3
Participant 2	Per	M	49	36	13	Y	100/Yr 5
Participant 3	Per	M	49	45	4	N	74/Yr 3
Participant 4	Per	F	49	42	7	Y	97/Yr 5
Participant 5	Per	M	58	36	22	N	84/Yr 3
Participant 6	Per	M	50	42	8	N	*62/Yr 2
Participant 7	Per	M	49	37	12	Y	*64/Yr 2
Participant 8	Per	M	53	24	29	Y	91/Yr 4
Participant 9	Per	M	61	48	13	N	98/Yr 4
Participant 10	Per	F	52	24	28	Y	89/Yr 4
Participant 11	Per	M	60	36	24	Y	85/Yr 3
Participant 12	Per	M	68	42	26	Y	107/Yr 4
Participant 13	Per	M	65	60	5	Y	79/Yr 2
Participant 14	Per	M	48	30	18	Y	78/Yr 3
Participant 15	Per	M	55	42	13	N	67/Yr 2
Participant 16	Per	F	59	36	23	Y	98/Yr 4
Participant 17	Per	M	48	24	24	N	86/Yr 4
Participant 18	Per	M	68	36	32	N	105/Yr 4
Participant 19	Per	M	63	48	15	N	105/Yr 4
Participant 20	Per	F	58	36	22	Y	94/Yr 4
Participant 21	Per	M	56	24	32	n/a	99/Yr 4
Participant 22	Rec	M	66	30	36	N	103/Yr 4
Participant 23	Rec	M	67	30	37	Y	115/Yr 5
Participant 24	Rec	F	49	24	25	Y	100/Yr 5
Participant 25	Rec	M	46	36	10	N	103/Yr 5
Participant 26	Rec	M	55	42	13	N	80/Yr 3
Participant 27	Rec	M	48	36	12	N	100/Yr 5
Participant 28	Rec	M	48	24	24	N	99/Yr 5
Participant 29	Rec	M	61	48	13	N	113/Yr 5
Participant 30	Rec	M	45	30	15	N	74/Yr 3
Participant 31	Rec	F	49	30	19	N	99/Yr 5
Participant 32	Rec	M	54	24	30	N	104/Yr 5
Participant 33	Rec	M	57	36	21	N	106/Yr 5
Participant 34	Rec	F	50	24	26	N	99/Yr 5
Participant 35	Rec	F	54	36	18	N	91/Yr 4
Participant 36	Rec	M	48	30	18	N	88/Yr 4
Participant 37	Rec	M	68	30	38	N	105/Yr 4
Participant 38	Rec	M	56	24	32	N	106/Yr 5
Participant 39	Rec	M	54	36	18	N	92/Yr 4
Participant 40	Rec	F	59	36	23	N	84/Yr 3
Participant 41	Rec	F	47	36	11	N	82/Yr 4
Participant 42	Rec	M	55	48	7	N	84/Yr 3
Participant 43	Rec	F	48	33	15	Y	85/Yr 4
Participant 44	Rec	M	57	46	11	N	94/Yr 4
Participant 45	Rec	F	58	46	12	N	95/Yr 4
Participant 46	Rec	F	58	48	10	N	94/Yr 4
Participant 47	Rec	M	41	36	5	N	75/Yr 3
Participant 48	Rec	M	55	22	33	Y	82/Yr 3
Participant 49	Rec	M	48	36	12	N	61/Yr 2
Participant 50	Rec	M	47	24	23	Y	85/Yr 4
Participant 51	Rec	M	61	42	19	N	98/Yr 4
Participant 52	Rec	F	52	36	16	Y	76/Yr 3

Note. All ages are reported in months. One participant's family history was not available (n/a). Two participants' status was confirmed via follow-up with the family (denoted by asterisks). Status Per = persist; Rec = recovered; F = female; M = male.

Appendix B

Model Accuracy for Expanded Cutoff Scores

Sensitivity, specificity, positive and negative predictive value by model using predicted probability of persistence cutoff of 0.30.

Model	Sensitivity	Specificity	Positive prediction	Negative prediction
BBTOP-PPI only	0.75	0.48	0.48	0.75
NRT score only	0.79	0.35	0.44	0.71
WSLD only	0.79	0.55	0.54	0.80
Family history only	0.60	0.81	0.67	0.76
NRT score multiple variable submodel	0.95	0.69	0.67	0.95
BBTOP-PPI multiple variable submodel	0.90	0.84	0.78	0.93
Full multiple variable model	0.95	0.83	0.78	0.96

Note. BBTOP-PPI = Bankson–Bernthal Test of Phonology, Phonological Process Inventory; NRT = nonword repetition test; WSLD = weighted stuttering-like disfluency.

Sensitivity, specificity, positive, and negative predictive value by model using predicted probability of persistence cutoff of 0.50.

Model	Sensitivity	Specificity	Positive prediction	Negative prediction
BBTOP-PPI only	0.45	0.71	0.50	0.67
NRT score only	0.37	0.86	0.64	0.68
WSLD only	0.53	0.83	0.67	0.73
Family history only	0.60	0.81	0.67	0.76
NRT score multiple variable submodel	0.74	0.79	0.70	0.82
BBTOP-PPI multiple variable submodel	0.80	0.84	0.76	0.87
Full multiple variable model	0.95	0.90	0.86	0.96

Note. BBTOP-PPI = Bankson–Bernthal Test of Phonology, Phonological Process Inventory; NRT = nonword repetition test; WSLD = weighted stuttering-like disfluency.

Sensitivity, specificity, positive, and negative predictive value by model using predicted probability of persistence cutoff of 0.60.

Model	Sensitivity	Specificity	Positive prediction	Negative prediction
BBTOP-PPI only	0.35	0.94	0.78	0.69
NRT score only	0.21	0.97	0.80	0.65
WSLD only	0.42	0.86	0.67	0.69
Family history only	0.60	0.81	0.67	0.76
NRT score multiple variable submodel	0.63	0.90	0.80	0.79
BBTOP-PPI multiple variable submodel	0.70	0.90	0.82	0.82
Full multiple variable model	0.79	0.93	0.88	0.87

Note. BBTOP-PPI = Bankson–Bernthal Test of Phonology, Phonological Process Inventory; NRT = nonword repetition test; WSLD = weighted stuttering-like disfluency.