

Children Evaluated for Speech and Language Concerns: One Year Outcomes

By

Tonia Nicole Davis

Dissertation

Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in

Hearing & Speech Sciences

August, 2016

Nashville, Tennessee

Approved:

Stephen Camarata, PhD, CCC-SLP (Chair)

Fred H. Bess, PhD, CCC-A

Michael de Riesthal, PhD, CCC-SLP

Warren Lambert, PhD

Copyright © 2016 by Tonia Nicole Davis

All Rights Reserved

ACKNOWLEDGEMENTS

This dissertation research would not have been possible without the support of any number of people. A few of these people are particularly salient:

Dissertation committee: Stephen Camarata, PhD, CCC-SLP (chair); Fred H. Bess, PhD, CCC-A; Warren Lambert, PhD; Michael de Riesthal, PhD, CCC-SLP

Doctoral cohort: George Castle, MS, CCC-SLP; Hope Lancaster, PhD (honorary member); Sterling Sheffield, PhD, CCC-A; Laura Wilson, PhD, CCC-SLP; Hatun Zengin-Bolatkale, MS, CCC-SLP

Clinical colleagues and supervisors: Charles Hausman, MS, CCC-SLP; Marcy Sipes, MS, CCC-SLP; Sue Hale, MS, CCC-SLP, and the entire staff of the Vanderbilt Bill Wilkerson Center

Statistical support from: Warren Lambert, PhD

And, of course,

The families who entrust us with their children

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	III
LIST OF TABLES	VI
LIST OF FIGURES	VII
Chapter	
1: INTRODUCTION	1
Introduction	1
Eligibility and Enrollment to Discharge	4
Correlates of Positive Outcome	9
The Current Study	16
Specific Aims and Hypotheses	17
2: METHODOLOGY	19
Participants	19
Database	20
Variables	21
Reliability	26
Analysis Strategies	27
Additional Methodological Considerations	29
3: RESULTS	31
Check for Normalcy	31
Initial Speech/Language Evaluation	31
Discharge and Outcome.....	34
Supplemental Analyses.....	37
4: DISCUSSION	41
Summary.....	41
Implications	41
Predictor Variables	43
Length of Treatment	47
Public Health Implications	49

Limitations of the Current Study	50
Future Studies: Prospective Data Analysis.....	53
Conclusion.....	55
REFERENCES	57
APPENDICES	87
Appendix A: Non-English Speaking Children	87
Appendix B: ICD-9 Codes Found in the Medical Charts	89

LIST OF TABLES

Table	Page
1. Specific aims, hypotheses, statistical procedures, and degrees of freedom.....	64
2. Reliability and missing data.	65
3. Variable types by time of administration and level of measurement.....	67
4. Kurtosis and skew for intake language variables.....	68
5. Referral to treatment based on initial evaluation variables.....	69
6. Odds of referral to treatment for each nominal variable.....	70
7. Enrollment in treatment for 154 children who were made eligible based on initial evaluation variables.....	71
8. Odds of enrollment in treatment for each nominal variable.	72
9. Discharge based on initial and follow up evaluation variables.	73
10. Odds of each nominal variable within discharge group.	75
11. Outcomes odds ratios for receptive language.	78
12. Concordance of clinician-administered measures and parent-reported measures.	79

LIST OF FIGURES

Figure	Page
1. Initial plan for process from evaluation to outcome.	80
2. Final plan for process from evaluation to outcome.	81
3. Inclusionary and exclusionary criteria for participation.	82
4. Hazard function for “Goals Met” group.	83
5. Hazard function for “Non-Attender” group.	84
6. Hazard function for “Discharge to Schools” group.	85
7. Prospective data analysis theoretical model.	86

Chapter 1: Introduction

Introduction

Implementation science addresses the extent to which interventions can fit within real-world public health and clinical service systems (Damschroder et al., 2009). In medicine, studies taking place within active service systems are often referred to as “effectiveness” studies. Effectiveness studies are contrasted with efficacy studies. Across allied health disciplines, tightly controlled “laboratory” conditions are often employed to study efficacy for an intervention under “ideal” conditions whereas effectiveness is examined in “real world” clinical contexts. In the field of speech-language pathology, efficacy studies are considered “gold-standard” research protocols, such as randomized controlled studies, which are conducted in clinical “laboratory” conditions with interventionists who meet stringent criteria in procedural fidelity (e.g., Camarata, Nelson & Camarata, 1994). Effectiveness studies, in contrast, take place in community settings such as schools or outpatient clinics, as delivered by the speech-language pathologists (SLPs) who are directly serving the patients and who may or may not meet the procedural fidelity specifications seen in efficacy studies. In addition, the children in effectiveness studies are not required to meet the stringent inclusionary and exclusionary criteria of a clinical laboratory study, although they will likely demonstrate similar profile characteristics and general diagnosis. Recent reviews have indicated that in pediatric speech-language pathology, only 2% of speech treatment studies and 19% of language treatment studies can be termed effectiveness studies (Baker & McLeod, 2011, Olswang & Bain, 2013).

Despite active promotion of systematic intervention research and review in the field of speech-language pathology (ASH Foundation Implementation Science Summit, March 2014), there continues to be a dearth of treatment research generally (Olswang & Bain, 2013) and especially of effectiveness treatment research (Olswang & Prelock, 2016). In fact, there is little research that even describes how a community clinical caseload might evolve over time, i.e., how children progress through treatment while receiving school-based or outpatient services.

For implementation studies, “best practice” in early intensive behavioral intervention (which includes but is not limited to speech and language) is defined in a 2009 systematic review as a minimum dosage of 12 hours per week and a *minimum duration of 12 months* (Howlin, Magiati & Charman, 2009). A review of the effectiveness and implementation science literature in speech and language treatment yielded no consensus on the duration of treatment that children typically receive in schools or clinics. What is clear is that the vast majority of children in treatment studies are also simultaneously receiving services outside of or in addition to those provided within the efficacy study (e.g., Roberts & Kaiser, 2015). Moreover, it is not at all unusual for children to receive treatment through the schools or clinic for more than one year, and some children with speech/language disorders may be eligible for treatment from birth to 21 (IDEA, 2004; IDEA – Part C, 2011). But, there are few “dose” studies evaluating optimal duration or intensity of intervention. Rather, intensity and duration may be driven by financial parameters (e.g., reimbursement) or service availability (e.g., school caseloads for SLPs). In contrast, efficacy studies are often conducted for a much shorter time period. Law, Garrett, and Nye (2004) indicated that 30% of speech or language treatment studies took place in less than eight weeks, while Baker and McLeod (2011) found the average duration of speech treatment to be under four months (M=3.75 months). Moreover, it would be highly unusual for “real world”

implementation studies of pediatric speech and language intervention to be conducted over such a short time period.

Although a well-established method in implementation science research is to conduct prospective effectiveness studies, which take place in real-world settings with children and clinicians who may not perfectly meet the narrow selection criteria often employed in short-term treatment efficacy studies, another widely employed method utilized in implementation sciences in medical research is to examine data from existing clinics and analyze that data retrospectively. *Retrospective data analysis* is the process by which statistical methods are applied to data collected for clinical (or other) purposes to address questions within the available data set. Probably the most famous recent example of retrospective data analysis derived from 1.26 million children to determine that there was no association between vaccination and autism symptomology (Taylor, Swerdfedger, & Eslick, 2014). Other methods of retrospective data analysis have also been used within speech language pathology and related allied health disciplines. For example, Watson and colleagues (2013) used retrospective video analysis to identify differences in joint attention in children with autism, developmental disabilities, and typically-developing children at 9-15 months, prior to subsequent diagnosis of autism.

Retrospective data analysis has a number of benefits for the field of speech-language pathology. First, it allows research to take place (a) among a larger number of children than are customarily enrolled in efficacy studies, (b) across a larger variety of disorder typologies, and (c) across a greater length of time than laboratory-based efficacy studies. Second, retrospective data analysis offers a view of a population receiving services, rather than a selected sample (e.g, only those who agree to participate in treatment efficacy research studies). Finally, retrospective data analysis helps us to understand what the clinical process looks like in community clinical

settings: the process by which a child is evaluated and made eligible for services, enrolls in services, and, eventually, is discharged from services.

Eligibility and Enrollment to Discharge

When a parent seeks speech and language services, a speech and language evaluation is first conducted in order to determine whether a child is eligible for services. If the child meets eligibility criteria, they will then be offered enrollment in treatment services, typically through either educational or healthcare organizations. The length of time during which the child will be enrolled in services is not predetermined. Some children will spend only a few months in treatment, whereas other children will remain in treatment for years.

Identification and eligibility for speech and language services: Overview of Early Intervention Services in the US. The purpose of speech and language intervention is, quite simply, to improve children’s speech and language. In theory, early treatment may allow some children (those with more mild disorders) to “catch up” to their peers before they reach school age, and may prevent those children with more severe disorders from falling farther behind and provide them with functional communication (Rosenbaum & Simon, 2016). Although it is an open question as to whether early treatment actually provides these benefits (Camarata, 2014; Warren et al., 2011), these services are provided in some form (educational or healthcare-based) in all fifty states under the rubric of “early intervention.”

To receive speech and language services, children must first be deemed eligible for services. For better or worse, eligibility is typically determined based on performance within a single evaluation. There may be times when a clinician conducts a follow-up evaluation or

continued monitoring. This may occur when a child scores low relative to norm-referenced standards, but does not meet strict eligibility criteria, or if a child scores low enough to be eligible but the clinician suspects performance is due to non-language factors, such as compliance with testing. However, it is far more common that eligibility will be determined based on a child's performance in a testing session that may last one to three hours, with a clinician that is meeting the child for the first time. Despite these potential limitations, performance on this initial evaluation is often used to determine eligibility criteria, and for offering enrollment in treatment services.

Specific eligibility criteria vary by state and by institution. In general, children must test "low enough" (perform poorly enough) on a set of standard speech and language measures. These measures may be given to the child directly by the clinician (for example, the Preschool Language Scale, 2011), or they may be structured interviews with parents and parent checklists (for example, the Receptive Expressive Emergent Language Test, 2003). Scores can be calculated and determinations made to establish eligibility. There are no federal guidelines and only rarely are there health insurance guidelines for determination of eligibility. When health insurance guidelines do exist, their rationale is rarely available to the public. Therefore, each state that funds Early Intervention under IDEA Part C (2004; 2011) develops their own guidelines. Medical institutions, such as hospital-based speech and language clinics, also can develop their own eligibility guidelines.

Overview of Tennessee Early Intervention Services. The State of Tennessee provides the following guidelines for Tennessee Early Intervention Services (TEIS): children under three may be tested across communicative (speech/language), cognitive, adaptive, social-emotional and motor domains. To be eligible for services, a child must test with greater than 25% delay in

two areas (for example, communicative and cognitive) or greater than 40% delay in one domain. As the observed “percent delay” will inevitably vary by measure, making decisions based solely on standardized measures may be variable and thus unclear to parents. Consider, for example, a 24-month-old child who receives the Receptive-Expressive Emergent Language Test (REEL-3, 2003) and scores a 66 standard score (age equivalency 14 months; 42% delay) and who during the same appointment receives the Preschool Language Scale (PLS-5, 2011) and scores a 74 standard score (age equivalency 15 months; 37% delay) (Bzoch, League, & Brown, 2003, Zimmerman et al., 2011)¹. If communication is the only concern, this child meets eligibility criteria under the first measure but not the second. In addition, test measures may not match developmental norms if a child has some relative strengths and some relative weaknesses. For example, a 24 month old child with excellent play and social skills, but zero words can score as high as a 91 standard score on the PLS-5, even though developmental norms would suggest this child should have greater than 200 words (Chapman, 2000; Miller, 1981)².

Rather than rely solely on standard measures, some institutions such as Vanderbilt University in Nashville, Tennessee rely both on standard measures and clinician judgment. That is, even under the most stringent test protocols, clinical judgment can be a factor for these children who test at borderline across different measures. Thus, standard measures and clinical

¹ Derived from the REEL-3 and PLS-5 manuals. Age equivalencies were derived from raw scores via manual tables (REEL Appendix D; PLS Appendix C.1). Percent delay was calculated using the formula $[(\text{chronological age} - \text{age equivalency}) / \text{chronological age} \times 100]$. Hypothetical standard scores were derived based on the maximum 95% confidence interval for each test; i.e., a standard score of 74 on the PLS has a 95% confidence interval of [66-86].

² Hypothetical child was credited for items such as: uses representational gesture, participates in play routine, initiates a turn-taking game, demonstrates joint attention. Hypothetical child was not credited for items such as: imitates a word, uses at least five words, names objects in photographs. Raw scores were calculated and standard scores derived using the manual tables in PLS Appendix A.

judgment are important components of the initial eligibility evaluation and final determination of eligibility is made by the clinician using testing and observations, and recommendations are made to the insurance company paying for services or to the public agency providing intervention. In this approach to determining eligibility, standard measures are augmented by clinical judgment. In the above example, at Vanderbilt Bill Wilkerson Center (VBWC), a child with “normal” test scores on play scales but zero words at 24 months can be made eligible for services based on clinical judgement. There currently are no data to compare children assessed and made eligible for services through TEIS and children who are assessed and made eligible for services through VBWC, as children will rarely be evaluated through both government (TEIS) and medical (VBWC) systems. For this study, we will consider eligibility through the system to which we have access, which, in this case, is VBWC.

When children are deemed eligible, they can then be offered enrollment in services. If children are eligible for intervention at VBWC, the following outcomes could occur: first, children could enroll in services at VBWC; second, children could enroll in services elsewhere; third, parents may choose not to enroll children in services at all (Figure 1). Children in the second and third groups are considered “lost to follow-up” for the purposes of this study. For the children who are enrolled in services at VBWC, however, we have a great deal of potentially useful information.

From enrollment in services to outcomes. After enrolling in services at VBWC, a child receives a recurring appointment with a speech language pathologist (SLP). The most common recurring appointment schedule is a weekly appointment of one hour for six months, but more and less frequent appointment schedules may occur. The clinician identifies goals, writes a treatment plan, and provides treatment for the child and education for the family. Re-assessment

occurs annually, but the clinician may provide additional testing at any point in time after enrollment. Additional time – adding another six months of scheduled appointments, for example – is prescribed by the clinician, and insurance specialists determine whether the family receives coverage for the extended services. VBWC has no predetermined amount of time that children will remain in services or criteria for length of time (or age) at which discharge is mandatory. Because “clinician judgment” is a contributing factor for entry into, and exit from services the process may appear subjective – and from a measurement standpoint, it most certainly is – this model nonetheless provides a significant range of data. Some children will enroll and be discharged from services within a matter of months, while other children remain enrolled at VBWC for years.

For some children, a possible outcome after enrollment in treatment is discharge from services due to meeting goals and/or normalizing (retesting within normal limits, i.e. between 85 and 115 on a standardized language measure). Not all children will meet this goal within one year, and not all children will ever meet this goal. For these children, the optimal outcome will be higher levels of communication than they “would have achieved” without early enrollment in services. This goal is not measurable at the child level, but we know, for example, that the average language outcome for a child with Down syndrome was higher in a 2011 meta-analysis (Naess, Lyster, Hulme, & Melby-Lervag) than in a 1960 study (Kolstoe) – before early intervention or much in terms of structured, evidence-based intervention was available. (Due to the nature of the study design and psychometrics, a direct comparison cannot be made.) For the full cohort of children under three who enroll in speech language services, however, we may assume that one possible outcome is discharge from services due to meeting goals or normalizing.

For those children who are enrolled in services at VBWC, we can examine enrollment outcomes at one year. Possible outcomes include discharge by one year of enrollment or continued service provision. Some children will remain in services one year after they enroll in services. Other children will have been discharged at some point prior to one year. Using discharge and time-to-discharge as markers of success, we can establish potential correlates – or predictors – of early discharge.

Correlates of Positive Outcome

Few studies examine full cohorts of children enrolled in speech and language services – for example, a study might remove children with known intellectual disabilities, or only include children who are diagnosed with autism spectrum disorders. However, looking across studies, several possible predictors of positive outcome have been identified. Four of these predictor variables are receptive language at an earlier time point, presence or extent of additional medical concerns, overall cognition, and access to resources.

Receptive language. Receptive language at intake is a well-established predictor of subsequent expressive language outcome. Although research studies tend to be population-specific, early receptive language is predictive of outcome in late-talking children (Fernald & Marchmann, 2012), children with autism (Thurm, Lord, Lee, & Newschaffer, 2007), and children with intellectual disabilities (Roberts et al, 2008). Receptive language is tested for every child who receives an initial speech and language evaluation at VBWC, as well as at all later evaluation time points. Thus this information is available in the electronic medical record.

An example child is “Adrian.” Adrian was first seen at VBWC for a speech/language evaluation when she was 25 months old. Her scores on the Preschool Language Scale (PLS-5) were 104 receptive and 79 expressive (a 25-point standard score gap). She spoke only 6 words during the evaluation. Speech/language therapy was recommended. The treating clinician focused on parent training and imitation based treatment. After four months, Adrian was discharged with scores on the PLS of 111 receptive and 97 expressive. Adrian continued to demonstrate a number of developmental articulation errors, so follow-up was recommended. At 37 months, Adrian was seen for a one-year follow up. At follow up, Adrian’s scores on the PLS-5 were 110 receptive and 102 expressive. Adrian’s mean length of utterance (MLU) was within normal limits, she was over 80% intelligible, and when administered a standardized articulation measure, her articulation score was within normal limits.

Despite expressive language well below normal limits (8th percentile), Adrian made rapid, steady progress and was successfully discharged from treatment due to normalizing (expressive language gain of 18 standard score points). Receptive language has been a positive indicator of both overall outcomes (Howlin, Mawhood, & Rutter, 2000; Mawhood, Howlin, & Rutter, 2000; Sutera et al., 2007) and response-to-treatment (Justice, Meier, & Walpole, 2005). In identifying correlates of outcome in existing medical records, we expect high receptive language scores to be associated with dismissal from treatment due to meeting goals, while low receptive language scores would be associated with retention in treatment.

Medical conditions. Children with additional medical conditions are routinely excluded from efficacy research studies because these concerns are expected to impede their response to treatment and/or to add heterogeneity. Premature birth, failure to thrive, or long hospital stays are “at risk” factors for speech and language disorders (Barre, Morgan, Doyle, & Anderson, 2011).

Similarly, children with chromosomal anomalies known to cause intellectual disabilities are at high risk for speech and language disorders (i.e., Roberts et al., 2008). Other common exclusionary criteria for treatment efficacy research studies include developmental or genetic disorders, hearing loss, risk for autism, and motor disorders (such as cerebral palsy, which may impede motor speech) (c.f., Spreckley & Boyd, 2009).

At VBWC, there are two possible sources to detect the presence of additional medical concerns in the health record. First, all parents or guardians fill out a form listing over twenty possible medical concerns (including “other”) that may affect speech/language development. Second, the electronic medical record itself contains all previous diagnoses. Although there is no standard way to determine the relative weight of different diagnoses, these children are nevertheless a part of a clinical caseload – and rarely included within a treatment efficacy research sample.

An example child is “Ethan.” Ethan was first seen at VBWC for a speech/language evaluation when he was 32 months old in preparation for discharge from TEIS at 36 months; he had received speech/language services, physical therapy, and occupational therapy through TEIS. Ethan also received nutrition through a g-tube and was followed by neurology. At the time of his initial evaluation at VBWC, Ethan had spent 87 lifetime days in the hospital.

At initial evaluation, Ethan produced only one word approximation: “nuh.” Ethan enrolled in speech/language services at VBWC, where he received individual speech treatment, as well as group alternative and augmentative communication (AAC) treatment, where he learned to use an iPad to communicate. By his one-year follow up, on the Clinical Evaluation of Language Fundamentals (CELF-P2) he scored a 98 receptive and a 59 expressive; on the Goldman-Fristoe Test of Articulation (GFTA-2), he scored a 65. He was able to produce the

phonemes /p b m n t d/ and /ʌ/ in CV and CVCV syllable shapes. Ethan remained in treatment and was transferred to a new treating clinician prior to kindergarten and continues to be seen. At 63 months, Ethan communicates using 3-5 word utterances and facilitates listener understanding using an iPad or by using initial letter cueing (Kentner & Miller, 2009) on a letter-board. He can produce all age-appropriate consonants and syllable shapes, and produces the vowels /i æ a o u ʌ/. He continues with speech and language services at VBWC and also sees an SLP at school. Ethan continues to battle medical concerns: he has now spent a total of 114 cumulative days in the hospital.

Ethan made slow but successful progress in treatment, possibly due in part to his many medical needs, which were known at the initial time point. It is reasonable to hypothesize that children with relatively fewer medical needs will progress more quickly through treatment and be discharged, while children with many medical needs would remain in treatment for at least one year of therapy.

Overall cognitive level. Cognitive level refers to a child's ability to access and interpret knowledge. Non-verbal cognitive abilities include attention, memory, and executive functioning, which are tested in young children primarily through parent report and play-based tasks (Alloway, Gathercole, Willis, and Adams, 2004). Cognitive ability is sometimes captured by proxy via medical conditions, e.g., a child with Down syndrome (a chromosomal abnormality) is likely to have an associated intellectual disability (Roberts et al., 2008). However, many children will not have a diagnosed medical condition with known cognitive delays before the age of three (cf. Katz & Lazcano-Ponce, 2008), so deficits in cognition may not be detected even in these cases. Few children will have received direct cognitive testing by the time they fail to meet

speech/language milestones and are seen for an initial evaluation. In fact, language is often the first indicator of concern (De Giacomo & Fombonne, 1998).

An example child is “Jason.” Jason was first seen at VBWC for a speech/language evaluation at 25 months after a well-child checkup that revealed possible language delays. He had reportedly met all motor, cognitive, and socio-emotional milestones; he appeared slightly delayed in communication and adaptive behavior. He received speech/language treatment for almost a year before other delays in cognition were identified. At age six, he was given an intelligence test in school and scored a 74 standard score, which places him at the fourth percentile. This information became available because his Individualized Education Plan (IEP) was scanned into the medical record.

Research studies routinely exclude children based on overall cognition (cf. Spreckley & Boyd, 2009) or control for nonverbal cognition (van der Schiut, Segers, van Balkom, & Verhoeven, 2011) in predicting outcomes. The inherent assumption is that children with lower cognition take longer to make gains in treatment (Davis, Lancaster, & Camarata, 2015). However, cognition scores from IQ testing are not routinely captured within the electronic medical record, and are rarely recorded within the speech/language evaluation. Occasionally developmental indices will be available (i.e., Bayley Scales of Infant Development). Measures of nonverbal intelligence (spatial reasoning), attention, memory, or executive functioning are usually not available for children under the age of three (cf. Swisher, Plante, & Lowell, 1994). In addition, many families may not wish to have the overall cognition of their children tested and recorded, so this variable may not be available within the electronic medical record.

Access to resources. Access to resources can take many forms. Socioeconomic status (SES) and parental educational achievement are often associated with better access to resources

(Lehane & McCarthy, 2007; Magunson & Duncan, 2005; Shavers, 2007), whether those resources include transportation, insurance status, ability to navigate the medical referral system, childcare or eldercare while attending treatment sessions, or other forms of support resources. Most of these variables do not exist within the medical record; for example, there would be no information on transportation (i.e., individual vehicle, public transportation system, medical transport, borrowed vehicle). Therefore, there are limitations on which resource variables can be addressed in the medical record.

A few variables are available in the medical record to estimate access to resources, including median income by zip code, insurance status, and distance from the VBWC clinic. Median income by zip code is a common variable when individual socioeconomic status is not available (Geronumous & Bound, 1988). Low-income zip codes are associated with non-adherence to medical treatment in adult populations (Fischer et al., 2011). Similarly, absence of insurance is associated with non-adherence in adults (Briesacher, Gurwitz, & Soumerai, 2007). Most studies of medical adherence and non-adherence are conducted in adults, so it is unclear whether these relationships exist in children. VBWC will bill government-assisted insurance such as TennCare, while more local facilities may not, which will likely affect the relationship between insurance status and treatment outcomes. In addition, we can calculate distance from VBWC by zip code, which may be an independent resource variable (i.e., enrollment in treatment) or a covariate. Families who need to travel extended distances to participate in services are likely to have additional barriers (e.g., car trouble) than families who live close by (cf. Strauss et al., 2006). In addition, families who live nearer VBWC may receive all of their medical care at VBWC, while families who travel extended distances are likely to have primary care providers closer to home; this may affect enrollment, continued attendance, and follow-up.

Compare the following two cases, which were seen at VBWC. “Leroy” was seen at VBWC for a speech/language evaluation when he was 22 months old. His parents, both physicians, were concerned because his older sister seemed to have progressed in speech and language more quickly than he. Leroy’s scores fell within low-normal limits on the PLS-5. Parents were counseled in language-facilitation techniques and follow-up evaluations were administered every six months. Now six years old, Leroy is being seen for a few residual articulation errors, but his parents are confident that his “lisp” will be corrected, and there are no academic concerns.

In contrast, “Lucy” was seen at VBWC for a speech/language evaluation when she was 23 months old. Her mother, too, was concerned about her based on her older children’s development. Lucy scored within low-normal limits on the PLS-5, but the clinician determined that she should be enrolled in services at VBWC – especially because her scores were too high to qualify through TEIS. However, Lucy’s mother was unable to bring her to VBWC for treatment. Lucy was seen for follow-up at VBWC when she was four years old. At this time, she scored below a quotient of 80 on all measures ($M=100$, $SD=15$ for norms). Based on her scores, she may have qualified to be seen through a preschool near her family’s home. Because she was never seen for treatment services at VBWC, and has not returned for another follow up evaluation, her outcomes are not available in the electronic health record.

At initial evaluation, Leroy and Lucy appear to be very similar: their ages and test scores are virtually identical, and their parents had similar concerns. The major difference is treatment recommendation, which is likely based on variables that are not directly accessible through the medical record. However, demographic variables related to access to resources may help to explain the difference in recommendation. Leroy’s family lived in an affluent part of Nashville

close to VBWC; they had excellent insurance. Lucy's mother lived over 45 minutes away from VBWC in a rural community with poor transportation options; Lucy's evaluations were provided through TennCare, Tennessee's Medicaid Program. We might expect that children from families with better access to resources would be more likely to receive treatment and to be dismissed from treatment, while children from families with less access to resources would either fail to enroll, fail to participate (i.e., dismissal for low attendance), or remain in treatment for longer. This hypothesis is consistent with literature on medically-matched children from high- and low-SES populations (Wild et al., 2013). Resources that can be calculated in this study include median income by zip code, presence or absence of insurance, and distance to VBWC.

Receptive language, presence of medical concerns, overall cognition, and access to resources are all potential predictors that may explain some portion of the variability in child outcome after enrollment in services. Receptive language, medical concerns, and access to resources are variables which are available within the electronic medical record; cognitive ability is not available.

The Current Study

The current study uses data derived from the electronic medical records of children who were seen at VBWC for initial evaluation of speech and language and whose eligibility was determined at that time. The benefits of using large clinical databases to develop characteristics and patterns of patients are well established. In medical studies, these databases often provide information on risk factors and treatment outcomes (cf. Beitchman et al., 2001). In addition, all children seen for evaluation are included in the study, rather than all children who voluntarily

enroll in a research study, a population that may be limited by demographic and personal variables. The inherent limitation of a database study is that some desirable variables may not be available (e.g., cognitive ability).

Ultimately, in future studies, the electronic medical records of children who enroll in services at VBWC may be able to provide valuable information about the effectiveness of specific treatment procedures on outcome. First, however, it is useful to describe the population of interest. Moreover, the population at VBWC is larger and more varied than what is in the literature and is representative of community-population based EI services. This first step in using medically-derived databases is to discover what children exist within the medical record, and what variables we might use to describe them.

Specific Aims and Hypotheses

The aims of the study are to describe children who are referred for speech/language evaluation at the Vanderbilt Bill Wilkerson Center over the course of a year at two time points, initial intake (and subsequent enrollment), and follow up and/or discharge. Hypotheses, statistical measures, and necessary degrees of freedom are found in Table 1. The expected flow of participants from intake to discharge is found in Figure 1. The observed flow of participants from intake to discharge is found in Figure 2.

Aim 1: To compare children who are determined to be eligible and children who are deemed to be ineligible based on their initial evaluation performance and to identify predictors of eligibility status. **Hypotheses:** Children who are evaluated and judged ineligible for services will (a) perform better on standard measures and (b) demonstrate relatively fewer medical concerns

than children who are evaluated and judged to be eligible services. That is, children with (a) lower language scores and (b) more medical concerns will be more likely to be deemed eligible for services.

Aim 2: To compare children who are enrolled in treatment at VBWC to children who do not enroll in treatment at VBWC after becoming eligible in their initial evaluation and to identify predictors of initial enrollment. **Hypotheses:** Children who enroll in treatment at VBWC will (a) live closer to VBWC as measured by distance by zip code and (b) have a relatively greater number of medical concerns than children who do not enroll in treatment at VBWC. Children who (a) live within 20 miles of VBWC and (b) have medical concerns will be more likely to enroll in treatment.

Aim 3: To describe groups of children who are discharged from treatment and those who remain in treatment at one-year follow-up and to identify predictors of discharge. **Hypotheses:** Children who are discharged from treatment due to meeting goals or reaching normal limits will display (a) higher receptive language, (b) fewer medical concerns, and (c) higher income by zip code as compared to children who remain in treatment. Children who (a) have receptive language within normal limits, (b) have no medical concerns, and (c) come from zip codes with incomes above 1.5 times the federal poverty line will be more likely to be discharged due to meeting goals or reaching normal limits.

Chapter 2: Methodology

The aims of this study are addressed using the following methods addressed in Table 1. Aim 1 compares children made eligible for treatment to children who are not made eligible for treatment and determines the odds that a child is made eligible for services given a specific predictor variable. Group comparisons are made using a series of t-tests for ordinal and interval variables. Odds are determined using odds ratios using nominal variables. Aim 2 compares children who enroll in treatment to children who do not enroll and determines the odds that a child is enrolled given a specific predictor variable. Group comparisons are made using a series of t-tests for ordinal and interval variables. Odds are determined using odds ratios using nominal variables. Aim 3 compares four groups of children: children who are discharged due to meeting their goals, children who remain in treatment, children who are discharged to the schools to continue treatment, and children who are discharged due to non-attendance. Group comparisons are made using ANOVA while odds are determined for each discharge group.

Participants

The study analyzes patient records from January 1, 2013 to December 31, 2013 to create a database of children who received speech-language evaluations. Children were included if they: (a) received a speech/language evaluation between the dates listed above, (b) were between 18 and 30 months at the time of evaluation, and (c) had not previously been seen for speech/language evaluation. Due to the difficulty in accurate measurement of language ability in non-English speaking or bilingual children, these children were excluded from the current analysis, and information on these cases are available in Appendix A: Non-English Speaking

Children. Demographic information was collected and compared to the city-wide demographics in which the evaluation occurred.

Database

Data extraction. The PI accessed medical records containing protected health information through Starpanel, a Vanderbilt-owned, secure medical record program. The PI obtained medical record numbers of children eligible for the study through the DHSS Research Subject Finder Request. The DHSS Research Subject Finder Request allows DHSS investigators to obtain electronic medical record numbers for children who are seen in the Vanderbilt Bill Wilkerson Center (VBWC). Children who were (a) seen for a speech/language evaluation between January 1, 2013 and December 31, 2013, and (b) no more than 42 months of age on December 31, 2013 were included.

After the PI obtained the medical record numbers of children who may have been eligible for the study (n = 532), the PI manually reviewed each medical record for the inclusionary and exclusionary criteria as described in Figure 3. All children who met inclusionary and exclusionary criteria were entered into the database (n = 198).

Data management. Data was derived from medical records. The variables of interest (defined below) were entered into RedCAP software (Harris et al., 2009) with protected health information (PHI) removed. The PI was responsible for transferring data from medical records to a REDCap database. The declassified information included: patients' age, family demographic factors, standardized and unstandardized information from clinical assessment, progress through therapy, outcome and discharge data. All dates were converted based on the

date of evaluation (e.g., age at evaluation, time passed since evaluation). A total of 166 variables were entered into the REDCap database, of which 25 are used for analysis. Percent missing data is found in Table 2.

After all information was transferred to REDCap, the physical copy of the medical record numbers was destroyed by shredding and disposed of per Vanderbilt Medical Center guidelines. The following information, which was reviewed in Starpanel medical records, was not entered into the declassified REDCap database: patient name, date of birth, date of assessment, address, medical or social security number, health insurance information other than status. This process was approved by the Vanderbilt Internal Review Board (IRB #151135).

Variables

All variables in the data set are found in Table 3. Variables are divided based on when they were administered and are classified as nominal, ordinal, or interval/ratio. In many cases, different variables were used for group comparisons and for odds ratios computation. For example, “receptive language” was defined as a standard score (interval/ratio) as well as a criterion cutoff for percent delay (nominal). The interval/ratio variable was used for group comparisons, but the criterion cutoff was used for odds ratios.

Eligibility, enrollment, and discharge. Children were determined to be eligible for services if the evaluating clinician recommended treatment of any amount (one visit a month to multiple visits a week). Children were not determined to be eligible for services if follow up was recommended (typically 3-6 months later) or if no recommendations were made. Children were determined to have enrolled in services if they attended *at least five treatment sessions*. The

purpose of this decision was that (a) the variable “placed on waitlist” did not seem a valid measure of enrollment, (b) for children who attended 1-2 sessions, it was not always clear if they were being seen for treatment or for follow up, and (c) the goal of the research question was to understand discrepancies between families who chose to enroll and families who did not choose to enroll. Families who attended for less than five treatment sessions were determined to non-enrollers rather than non-attenders.

Discharge was initially quantified in two ways: presence or absence of discharge (i.e., remained in treatment), and time to discharge in days. However, over the course of data collection, it became clear that many children were discharged not due to meeting goals or reaching normal limits, but due to participation (non-attendance) or external factors (i.e., the family moved to a different state). In addition, a large number of children turned three over the course of the one-year follow up, at which point they may have become eligible for services through the school system. Therefore, the final discharge variables were: remain in treatment, discharge due to goals met, discharge due to non-attendance, discharge to schools, and discharge – other (see Figure 1, Figure 2). Seven children were discharged for “other” reasons and are not included in the analysis. Of these seven children, two were discharged due to moving to other states, one was discharged because the child returned to the biological home from foster care, and four children were discharged “at parent request.” No outcome data is available for any of these children.

Demographic variables. The following demographic variables were considered: age at initial evaluation, minority status, gender, median household income by zip code, insurance status, distance from VBWC, and presence of non-English languages at home (which could affect scores on English-standardized measures). *Age at initial evaluation* was defined in months

(e.g., 24 months). *Minority status* was defined as non-Caucasian non-Hispanic race and ethnicity, with children whose race and ethnic information was defined as “unknown” or “declined to answer” presented as missing data. *Gender* was defined as male or female. *Presence of non-English languages* in the home was established as a nominal variables where greater than 25% of the child’s interactions took place in a language other than English per parent report. For example, if a parent reported the home environment was “mostly Spanish,” the child would be considered a bilingual speaker/English language learner for the purpose of this study. However, if a parent reported that “Grandmother speaks Arabic, but child only sees grandmother once a week,” and otherwise spoke English at home and at daycare, the child would be considered an English speaker. The purpose of this variable was to identify cases in which English-normed language protocols would not be appropriate or valid for the child in question.

Median income by zip code was defined using the patient address to determine zip code and information from the United States Census Bureau 2013 data. Per capita income was used. In addition, to create a nominal variable, 1.5x the federal poverty line was established per United States Department of Health and Human Services 2013 guidelines. In 2013, 1.5x the federal poverty line was an income of \$23,265. This income level represents per capita poverty, not poverty by zip code. 1.5x the federal poverty line was chosen as a cutoff due to research from the National Center for Children in Poverty (NCCP, 2014), which identifies “true” poverty as between 1.5 and 3 times the federal poverty line.

Insurance status was initially defined as no insurance, public insurance (TennCare), or private insurance; however, no children were without insurance in the data set. 48% of children in the data set were seen through public insurance. TennCare covers children of families within 2.5 times the federal poverty line. In addition, children who are seen through TEIS receive

TennCare (often as secondary coverage).

Distance from VBWC was defined using the zip code (not the full address) of the patient and the VBWC physical address. This information was entered into Google Maps. Because the route determined changed depending on the time of day data was entered, discrepancies are discussed in the Reliability section below. For the nominal variable, 20 miles was set as the criterion cutoff (range <1 mile to 92 miles). 20 miles was chosen based on data that suggested people are willing to drive a maximum of 17-22 miles for regular services (Durning, 1992).

Medical variables. Medical history is not typically used within speech and language research except as an exclusionary criterion, so there was little previous research to guide the consideration of medical history. On the one hand, a nominal or ordinal variable reduces variability dramatically: a child with a history of ear infections is “considered” equal to a child with a history of profound sensorineural hearing loss using a nominal classification for presence or absence of hearing loss in the record. However, due to the lack of data in the matter, weighing medical events (profound sensorineural hearing loss weighted higher than ear infections) interval/ratio classification was likely to be a highly unreliable method. Therefore, despite the inherent limitation, nominal and ordinal variables were created.

Child medical history could be found through combing the medical record as well as via parent report on the intake evaluation forms. These two methods were combined with duplicates removed to identify an ordinal variable: *number of medical history events*. In addition, two nominal variables were created. The nominal variable was presence or absence of medical history. However, few children (16%) reported no medical history at all. Therefore, a secondary nominal variable was created as medical history less than four events. This variable was determined because the average medical history was less than two and the range was 0-12. The

same method was used to derive an additional variable, *family history of speech/language disorders*.

Speech/language variables. The following speech/language variables were considered: standard and raw scores on language measures administered by a clinician, standard and raw scores on parent report measures, and, if applicable, standard and raw scores on articulation measures administered by a clinician. In addition, the following parent-reported variables were considered: number of different words produced by the child and estimated percent intelligibility. These variables are part of the intake questionnaire given to parents. At one-year follow-up, only standard and raw scores on language and articulation measures were considered.

We expected a wide range of standardized measures. However, as it turned out, 96% of clinician-administered measures used the Preschool Language Scale-5 (2011) at intake (at outcome, the PLS-5 accounted for 20% of administrations; the CELF-P2 (2004), the Oral and Written Language Scale (OWLS-2, 2011), The Peabody Picture Vocabulary Test (PPVT-4, 2007) and Expressive Vocabulary Test (EVT-2, 2007), the Fluharty Preschool Speech and Language Screener (2000), and New Reynell Developmental Language Scales (2011) accounted for the additional 80%). No children received two clinician-administered measures of language unless one was receptive and one was expressive (i.e., the PPVT for receptive vocabulary and EVT for expressive vocabulary); in these cases, each test was entered independently, but no omnibus score could be entered. For parent-reported measures, common measures included the REEL-3 (2003; 64%), the Developmental Assessment of Young Children (DAYC-2, 2012), subtests of the Vineland Adaptive Behavior Scales (2005) or the Mullen Scales of Early Learning (1995), and the Rossetti Infant-Toddler Language Scale (2006; does not yield standard scores). There were no instances where children received two standardized parent-reported

measures. When children received both a standardized parent-reported measure and a criterion-reference parent-reported measure (i.e., the REEL and the Rossetti), the standardized measure was used for analysis.

To create a nominal variable, percent delay according to TEIS guidelines was used. Two nominal variables were created: 25% delay and 40% delay.

Reliability

A secondary coder entered 25% of child medical records to determine reliability. Data entry reliability was determined as percent agreement. Although percent agreement is not an ideal reliability measurement for nominal data due to high levels of chance agreement, it is an acceptable measure of agreement for ordinal, interval, and ratio data (Hartmann, 1977), which constitutes the majority of variables. 161 (97%) variables were entered with percent agreement at or above 95%, which suggests accurate data entry across most variables. The three variables below 95% agreement were: parent report number of different words (85%), parent reported percent intelligibility (85%), and distance from Vanderbilt by zip code (50%). Upon investigation, it was revealed that all disagreements in parent report number of words and parent report percent intelligibility were due to missing data, i.e., one coder found this information in the medical record and the second coder did not. For entries where both coders found the information, agreement was 100%. Therefore, data entry accuracy is not a concern, but missing data may be more prevalent on these particular variables than on other variables in the sample.

Distance from VBWC had an initial percent agreement of 50.6%, which is cause for concern. However, further investigation revealed that all disagreements were less than two miles

apart ($M = 0.2$ miles) and likely represent discrepancies in Google Maps as measured on different days and different times of day (i.e., different routes suggested). This variable was recalculated after rounding to the nearest mile, at which point percent agreement between the initial coder and the reliability coder was 92%.

An intraclass correlation coefficient (ICC, Shrout & Fleiss, 1979) was used for two additional variables: patient medical history and family medical history. ICC was used for these variables because they were derived from the medical record as ordinal variables. For patient medical history, $ICC=.701$ and for family medical history, $ICC=.614$. ICCs above .6 are considered “good” and ICCs above .75 are considered “excellent” (Cicchetti, 1994; Hallgren, 2012). Therefore, both of these variables may be considered to have good reliability.

A total of 166 variables were coded for each child. All derived variables are included in Table 2 with reliability and percent missing data included. Derived variables include all variables that were not directly copied from the medical record, including simple calculation derivations such as for age and time-to-discharge. In addition, all variables with reliability below 95% are included in the table. Therefore, the table can be considered a conservative description of reliability.

Analysis Strategies

For each aim, the analysis strategy is described below.

Aim 1: To compare children who are determined to be eligible and children who are deemed to be ineligible based on their initial evaluation performance and to identify predictors of eligibility. This was tested through a series of t-tests and through odds ratios. T-tests were

chosen as a measure of group differences when there were two groups and when the data was distributed normally. (Mann-Whitney U was chosen *a priori* as a non-parametric statistic, had the data not met criteria for normalcy). Odds ratios were chosen as a measure of association between a categorical variable and an outcome. The purpose of using both statistics was to analyze both categorical (nominal/ordinal) and continuous (interval/ratio) variables, and, when appropriate, to compare the results. This section applies traditional criteria (Cohen, 1992), such as $p(\alpha) < .05$ and power = .80 to describe the study's sensitivity to results (Kraemer, Mintz, Noda, Tinklenberg, & Yesavage, 2006). With approximately 200 participants in the sample, a t-test of means using Pass Power software (Hintze, 2008) found the ability to detect differences of Cohen's $d = 0.35$ or larger.

Aim 2: To compare children who are enrolled in treatment at VBWC to children who do not enroll in treatment at VBWC after becoming eligible in their initial evaluation and to identify predictors of enrollment. This was tested through a series of t-tests and through odds ratios. With approximately 150 participants in the sample, a t-test of means found the ability to detect differences of Cohen's $d = .32$ or larger.

Aim 3: To describe groups of children who are discharged from treatment and those who remain in treatment at one-year follow-up and to identify predictors of discharge. This was tested through ANOVA and odds ratios. Originally, Aim 3 anticipated only two groups and t-tests were planned. However, when four groups were chosen to be analyzed, ANOVA was chosen as a new statistic. ANOVA is appropriate in situations when more than two groups are compared across the same dependent variable. The odds ratio analysis, which was completed within groups rather than between groups, remains the same as in Aims 1 and 2. With

approximately 70 participants in the sample, an effect size of $d=.51$, or $\eta^2 = .06$ or above can be detected (transformation calculated from Cohen, 1988; Rosenthal, 1994, p. 239).

Additional Methodological Considerations

The following approaches were applied for all analyses:

Multiple imputation. In a retrospective chart review, some missing data is to be expected. This can arise from loss to follow up, missing data entry or other factors. The amount of missing data for each variable is presented in Table 2. A widely used approach is to employ *multiple imputation* to estimate the missing data point. That is, multiple imputation corrects for missing data by averaging the outcomes across multiple estimations (Sterne et al., 2009). Conceptually, the sample data are utilized to create an “expected value” for a missing data point, which is then entered in the data base. Although single imputation can be used, multiple imputation which includes more than one estimate of the missing score is often more accurate. Multiple imputation was applied to the data base herein. Imputation increased the likelihood of statistical significance when group comparisons were made on this data. Because all participants could be included, which, at the least, increased the effective power of the analyses. Demographic variables such as race and ethnicity, as well as all standard score variables were imputed using multiple imputation and the hypotheses were run on both the original data sets and the imputed data sets. Imputation changed the results of one variable: clinician-administered expressive language standard scores.

Multiple tests and familywise error. Although the hypotheses in Chapter 1 include only seven direct statistical tests ($n=198$), a larger number of statistical comparisons were made

over the course of the study. The likelihood of familywise error increases across the number of hypotheses. *Familywise error* refers to the probability of Type 1 error (false positives) in multiple hypothesis testing (Tabachnik & Fidell, 2001). Adjusting for familywise error across nominal, ordinal, and ratio variables is not trivial nor straightforward, particularly in an exploratory-descriptive study wherein a priori effect sizes are unknown.

One method for correcting for familywise error is to adjust the probability value of Type 1 error. The Sidak correction (Sidak, 1967, Tabachnick & Fidell, 1996) is one method to adjust the p-value for multiple t-tests. Using the Sidak method, each hypothesis is tested at the level of $1-(1-.05)^{1/m}$ where m is the number of tests planned. Using this method, only tests with p-values less than .017 should be considered significant (rather than p-values less than .05). Significance is therefore reported at the level of $p < .05$ or $p < .01$ and can be interpreted by the reader on the basis of the Sidak correction.

Concordance between group comparisons and odds ratios. Because any Type 1 error is a false positive, another consideration is concordance between the group comparisons and the odds ratios at each level of hypothesis testing. When there is concordance between the two methods of analysis, the likelihood of a false positive is reduced. Of course, there may be circumstances where the odds ratios, which use nominal data and therefore condensed variability, may turn out significant where the group comparisons, which use a full range of variability, do not. However, where there is concordance, we may be more confident that the results are significant *within this particular sample*, and are unlikely to be attributed to Type 1 error.

Chapter 3: Results

The results of the analyses are provided below. Broadly, two primary analysis approaches were employed. The first is direct analysis of the predictors, along with traditional tests of statistical significance. Second, odds ratio analyses are presented. This latter approach was included because, from a clinical perspective, the “relative risk” approach can be applied to “real world” clinical decision making, which is nominal in nature (e.g., eligible or ineligible for services). Raw data are presented for means and standard deviations, but imputed data are provided for all statistical tests as well.

Check for Normalcy

The four main language scores were used to check for normalcy at the whole sample level and the treatment level. The sample was considered normal if skew was between -1 and 1 and kurtosis was between -3 and 3 (Tabachnick & Fidell, 2001). At the level of the whole sample, all language variables met this test of normal distribution. At the level of the treatment group, all language variables were also normally distributed. See Table 4 for details.

Initial Speech/Language Evaluation

Group differences were considered at the level of referral to treatment and enrollment in treatment, based on initial speech/language evaluation results. In addition, group differences in demographic information was considered. Because the sample met the criteria for homogeneity of variance (Table 4), t-tests are used to determine group differences for interval variables.

Mann-Whitney U, a non-parametric statistic, is used to determine group differences for ordinal variables.

Referral to treatment. Of 198 English-speaking children in the database, 154 (79%) were deemed eligible for treatment and 39 (20%) were not referred for treatment. Two children were not referred to treatment at their initial assessment but were referred to treatment one year later; these children cannot be analyzed separately due to low sample size. They are included in the non-referral group in this analysis. We hypothesized that children who are evaluated and judged ineligible for services will (a) perform better on standard measures and (b) demonstrate fewer medical concerns than children who are evaluated and judged eligible for services.

Children in the “eligible” group demonstrated lower receptive and expressive language than children not referred to treatment (Table 5). Eligible children demonstrate receptive language approximately 1.5 standard deviations below the mean (clinician-administered $M = 77.80$, $SD = 15.34$; parent-reported $M = 79.36$, $SD = 16.96$). In contrast, children judged “ineligible” for treatment demonstrate receptive language within normal limits (clinician-administered $M = 101.44$, $SD = 11.40$; parent-reported $M = 94.58$, $SD = 14.31$). For expressive language, discrepancies between clinician-administered measures and parent reports were found; these are discussed in post-hoc analyses. Children eligible for treatment demonstrate expressive language of 1.5-2 standard deviations below the mean (clinician-administered $M = 78.60$, $SD = 10.06$; parent-reported $M = 68.22$, $SD = 1.35$). In contrast, children deemed ineligible demonstrate expressive language from within normal limits to one standard deviation below normal limits (clinician-administered $M = 101.19$, $SD = 8.67$; parent-reported $M = 85.52$, $SD = 8.99$). Parent reported number of different words also differed, with eligible children having an average of 7 words and ineligible children having an average of 23 words. Only half ($n=16$) of

the children in the ineligible group had measures of parent-reported number of different words, likely due to robust vocabularies that parents did not estimate at the time of evaluation.

Demographic variables were analyzed additionally using odds ratios (Table 6). Rather than detecting group differences, odds ratios determine the likelihood that children will be referred to treatment based on demographic group. Children from minority backgrounds had a greater chance of referred to treatment [OR 1.350 CI 1.064 - 1.712] than children from Caucasian backgrounds, despite the lack of group differences between the referral to treatment and no-treatment groups. In addition, children with a history of premature birth had a greater chance of being referred to treatment [OR 1.393 CI 1.006 – 1.927]. Children with a history of otitis media (middle ear infections) and children with a family history of language or learning disabilities were less likely to be eligible for treatment.

Enrollment in treatment. We hypothesized that children who enrolled in treatment at Vanderbilt will (a) live closer to Vanderbilt as measured by distance by zip code and (b) have greater medical concerns than children who do not enroll in treatment at Vanderbilt. The results indicated that 67 children were enrolled at Vanderbilt and 87 children were lost to follow up after initial evaluation. Based on intake scores, these children differed on distance from the clinic ($t=2.54$, $p<.05$, Table 7). As might be expected, children who did not enroll at Vanderbilt had a mean distance-from-Vanderbilt that was over 10 miles farther than children who did enroll at Vanderbilt. In addition, children who enrolled in treatment at Vanderbilt had lower expressive language scores on clinician-reported measures ($t=2.48$, $p<.05$), although not on parent-reported measures.

Demographic variables were again analyzed using odds ratios (Table 8). Children from families below 1.5 times the federal poverty were more likely to enroll in treatment [OR 1.316

CI 1.024 -1.692]. Additionally, children who were already receiving additional services were more likely to enroll in treatment [OR 1.342 CI 1.159 – 1.555]. Children who came from zip codes greater than 20 miles away and children with a history of otitis media (middle ear infections) were less likely to enroll in treatment.

Discharge and Outcome

Sixty-seven children were available for analysis of discharge and outcome. We hypothesized that children who are discharged from treatment due to meeting goals or reaching normal limits will differ from children who remain in treatment based on having (a) higher receptive language, (b) fewer medical concerns, and (c) higher income by zip code. During the final analyses, we used four different outcome groups: children who were discharged due to meeting goals (expected), children who remained in treatment (expected), children who were discharged due to non-attendance (expected, but no analyses planned), and children who were discharged to the school system. An additional seven children were discharged for other reasons and are described in Chapter 2, but were not included in discharge analyses. Means and standard deviations are available in Table 9.

Outcome groups. Each of the three variables of interest were significant across outcome groups. Receptive language was significantly different across outcome groups, $F(4,62)=34.121$, $p<.01$, $\eta^2 = .093$. Specifically, children who were discharged due to meeting their goals or reaching normal limits differed from the whole sample by having higher initial receptive language ($M = 94.83$ $SD = 8.61$) and children who remained in treatment differed from the sample by having lower initial receptive language ($M = 63.88$ $SD = 12.35$). Median income by

zip code was significant across outcome groups, $F(4,54)=3.989$, $p<.05$, $\eta^2 = .017$. Specifically, children who were discharged due to meeting their goals differed from the full sample by living in higher income zip codes ($M = \$49,622$ $SD = \$19,067$). Finally, patient medical history was significant across outcome groups, $F(4,52)=6.269$, $p<.05$, $\eta^2 = .063$. Specifically, children who were discharged due to non-attendance were less likely to have numerous medical history events ($M = .90$ $SD = 1.29$).

Further analysis revealed two additional variables that were significant for children discharged due to non-attendance. First, children were less likely to have received additional services (9% vs. 45%), which is in keeping with limited medical history events. Second, children discharged due to non-attendance were more likely to have a family history of language or learning disabilities ($F(4,52)=4.310$, $p<.05$).

Odds ratios are used to measure the relationship between a predictor variable and an outcome. Odds ratios may be particularly useful in describing nominal information (event occurrence), which cannot be interpreted using a correlation. Odds were calculated for each of the four outcome scenarios (Table 10). Children in the “Goals Met” outcome group had decreased odds of coming from families with incomes within 1.5 times the federal poverty line [OR .719 CI .579-.893], were less likely to have four or more medical concerns [OR .775 CI .656-.916], and were less likely to have receptive language delays of 25% or greater [OR .632 CI .448-.890]. Children in the “Remain in Treatment” outcome group had increased odds of receptive language delays of at least 25% [OR 2.684 CI 1.240-5.810]. Children in the “Non-Attender” outcome group had increased odds of coming from families near the federal poverty line [OR 1.458 CI 1.045-2.250], decreased odds of receiving any additional services (OT, PT,

TEIS, etc) at the time they sought evaluation [OR .743 CI .569-.970], and decreased odds of having a receptive language delay of 40% or more [OR .826 CI .685-.996].

Performance on standard measures at outcome. Children in the “Goals Met” group demonstrated outcome scores within normal limits for both expressive (M = 98.57 SD = 12.91) and receptive (M = 103.56 SD = 13.36) language, while children in the “Remain in Treatment” (M = 78.23 SD = 19.41 and M = 76.46 SD = 21.66, respectively) and “Discharge to Schools” (M = 80.88 SD = 7.55 and M = 75.50 SD = 12.76, respectively) groups demonstrated continued delays in language after a year of treatment. A comparable finding was found within the odds ratio analysis. Concurrent odds ratios were calculated for those outcome groups that had follow-up scores (Table 11). Children in the “Goals Met” group had decreased odds of any receptive language delay at outcome (no children demonstrated a receptive language delay of 25% or more in this group), while children in the “Remain in Treatment” and “Discharged to Schools” groups had increased odds of continuing receptive language delays.

Analysis by outcome scores. Ultimately, the decision to discharge is based on outcome factors (including outcome scores) and it is important to explore the relationship between intake factors and discharge. Therefore, two analyses addressed the effect of intake variables on outcome scores. First, a linear regression was used to predict the extent to which the three main variables of interest accounted for outcome language. Second, odds ratios for group membership based on outcome receptive and expressive language were computed.

A linear regression was completed using the three intake variables of interest (receptive language, median income by zip code, patient medical history) on the outcome variable of omnibus language. 56 children were used in the regression (67 in the total sample less the 11 non-attenders, who do not have outcome scores). The regression equation was significant at

$F(3,44) = 11.867, p < .01$. Adjusted R squared was .204, which accounts for 20% of the variance. Standardized Beta weights are reported as follows: for receptive language at intake, $\beta = .412, t = 5.176, p < .01$; for median income by zip code, $\beta = -.025, t = -.309, n.s.$; for patient history $\beta = -.184, t = -2.279, p < .05$. Although these regression equations are significant, they account for relatively little variance across a year.

Supplemental Analyses

Three supplemental analyses were run *post-hoc*. The first analysis is survival analysis. Survival analysis (Efron, 1988, Klein & Moeschberger, 2005) is a measure of “time-to-event,” where “event” was considered discharge. Children in the “Goals Met,” “Non-Attender,” and “Discharge to Schools” outcome groups were analyzed. Second, several subgroups are described *post-hoc*. These subgroups include children with hearing loss, children with a history of otitis media (middle ear infections), and children with a history of premature birth. Third, concordance between clinician-administered measures and standardized parent-reported measures is analyzed.

Survival analysis. “Time to discharge” is plotted in Figure 4, Figure 5, and Figure 6 for the three outcome groups that were discharged. Children in the “Goals Met” group demonstrated mean days to discharge of 325.011 [CI 315.216-334.807]. Although a few children were discharged around four months, the majority of children were discharged after eight months of treatment. Children in the “Non-Attendees” group demonstrated mean days to discharge of 294.998 [CI 202.766-344.125]. There are no patterns to time-to-discharge, with children discharged between three and ten months in a fairly linear pattern (remember that children who

received fewer than five sessions, who would have been discharged between 1-2 months, are not included in the analysis). Children in the “Discharge to Schools” group demonstrated mean days to discharge of 343.481 [CI 336.169-350.793]. Children in this group were typically discharged after eight months, which is ecologically valid: children cannot be discharged to the schools before 36 months (the entry age for school services). The maximum age of the sample was 30 months, so children would not have been eligible for discharge to the schools until after six months of treatment or more.

Subgroup descriptions. Subgroup analyses were not planned, but certain subgroups were identified during the analysis process. These subgroups include children with established hearing loss, children with specific medical concerns including a history of otitis media and a history of premature birth, and children from English Language Learner families.

Children with hearing loss. We did not anticipate many children with hearing loss within the sample of children receiving an initial speech/language evaluation, since the age range was 18-30 months and the Early Hearing Detection and Intervention (EHDI; Joint Commission of Infant Hearing, 2007) 1-3-6 guidelines suggest that these children should be identified well before 18 months. Audiological evaluation is a pre-requisite of speech evaluation in this age range, and 98% of children had an audiogram in their medical records. We found ten children (2% of the sample) who were diagnosed with conductive or sensorineural hearing loss during their pre-speech audiological evaluations, which were conducted by a licensed audiologist.

Of the ten children identified with hearing loss during pre-speech audiological evaluations, only one child had been previously, definitively identified with persistent hearing loss. However, nine out of 10 demonstrated risk factors (Meyer et al., 1999), and all nine had been referred for further audiological evaluation previously. Specifically, eight of these children

were placed in the Neonatal Intensive Care Unit (NICU) as infants; six of them had a history of premature birth; and four of them had failed newborn hearing screening in at least one ear. Likely due to distance (average distance to VBWC was 32 miles), only two of these children were seen for speech/language treatment at VBWC, both of whom remained in treatment at one year post evaluation. The majority continued to receive audiologic care at VBWC.³

Children with a history of otitis media. A high number of the children with a history of otitis media (n=78, 39% of the sample) were seen for initial speech/language evaluations. Children with a history of otitis media were less likely to be considered eligible for treatment [OR .721 CI .582-.893] and less likely to enroll in treatment [CI .559 CI .437-.715]. Children with a history of otitis media were no more likely to end up in any particular outcome group.

Children with a history of premature birth. There was also a large number of children with a history of premature birth (n=36, 18%). Children with a history of premature birth were more likely to be eligible for treatment [OR 1.393 CI 1.006-1.927] and more likely to belong to the “Remain in Treatment” group [OR 1.422 CI 1.137-2.012]. Children with a history of premature birth were also more likely to have at least four medical concerns; 69% of children with a history of premature birth had at least four medical concerns.

Children from families that did not speak English in the home. Finally, children with English Language Learner families were not included in the analyses for the 198 monolingual children analyzed in this data set. However, 58 such children were in the electronic medical

³ The remaining child had passed a newborn hearing screening (OAE) and an earlier audiological evaluation (OAE). After her pre-speech audiological evaluation, she was referred for an Auditory Brainstem Response (ABR) test and diagnosed with sensorineural hearing loss. Despite attempts to follow up by VBWC audiologists, speech-language pathologists, and social workers, she did not return to VBWC after her diagnosis.

record (23% of the total sample of children 18-30 months) and were coded according to the same procedures and included separately in Appendix A: Non-English Speaking Children.

Concordance of parent report and clinician administered report. 21 children received both a parent-reported measure and a clinician-administered measure (Table 12). Receptive language had high concordance between measures ($ICC = .870, p < .001$), while expressive language had relatively low concordance ($ICC = .509, n.s.$). Children who received both measures scored significantly higher ($p < .05$) on clinician-administered measures than on parent reported measures. One possible interpretation of these results is that clinician-administered measures of expressive language overestimate children's expressive language abilities. Because 96% of clinician administrations used the Preschool Language Scale – 5th edition, this interpretation is restricted to the PLS-5.

In addition, children who received both parent and clinician administered measures scored lower on clinician-administered expressive language than children who only received a clinician-administered test ($p < .05$). Children who received only parent-reported expressive language measures did not differ significantly from children who received both types of measures. These findings suggest that children who scored highly on clinician-administered measures of expressive language were less likely to receive a parent-reported measure in addition.

Chapter 4: Discussion

Summary

The 198 children's enrollment-to-outcome status from speech/language evaluation to one-year outcome was analyzed using retrospective data analysis from their electronic medical records. 154 children were deemed eligible for services (79%) and 44 were deemed ineligible for services at the time of evaluation. 67 of the eligible children were enrolled in treatment at Vanderbilt (34% total sample; 44% referred sample). Nine children were discharged from treatment due to meeting goals or reaching normal limits (13% of the treatment sample). An additional 40 children remained in treatment either at Vanderbilt or through the local school system (60% treatment sample). In addition, attrition from treatment as defined by discharged due to non-attendance accounted for 16% of the sample.

Implications

Perhaps the most important result relates to which children ultimately are eligible for intervention, enroll in treatment and then are discharged within one year. Within the Goals Met subgroup, a child was less likely to come from a home near the poverty line, less likely to exhibit significant medical concerns, and less likely to have a receptive language delay of 25% or greater. Within the Discharge to Schools group, no factors were identified. Within the Remain in Treatment group, a child was more likely to exhibit both a 25% and a 40% receptive language delay. Within the Discharge due to Non-Attendance group, a child was more likely to come

from a home near the poverty line, less likely to be receiving additional services, and less likely to exhibit a receptive language delay of 40% or more.

Put another way, children who come from middle class homes, have no additional medical concerns, and do not have significant receptive language delays are likely to have completed treatment within a year. Children with receptive language delays of 40% or more are likely to continue treatment beyond one year. And, children who come from families with incomes near the poverty line and who exhibit delays of less than 40% are more likely to become non-attenders than other children.

Based on intake factors, it is unclear why children are discharged to the school system. They do not differ from the rest of the children who received treatment on demographic factors (age, race, income, insurance status) and they also do not differ from the rest of the children based on percent delay at intake. Outcome factors may help to reveal this discrepancy. 100% of children who were discharged due to meeting goals demonstrated receptive language within normal limits at outcome (0% of children demonstrated a 25% receptive language delay or greater); 50% of children who remained in treatment demonstrated receptive language delays of 40% or greater. Among children discharged to the schools, 25% demonstrated receptive language within normal limits, 37.5% demonstrated receptive language delays between 25% and 40%, and 37.5% demonstrated receptive language delays of greater than 40% (note that data is missing for two children in this group).

Predictor Variables

Three hypothesized predictors of outcome were receptive language, medical history, and median income by zip code. As described in the introduction, I hypothesized that receptive language would be a positive-protective factor because relatively greater listening comprehension could relate to a child's ability to process and acquire new linguistic forms from the ambient environment. Poor receptive language would potentially reduce these incidental opportunities and perhaps require more direct support to learn new speech and language skills. As a related hypothesis, medical history could be related to increased inhibitory factors associated with relatively poorer health. Finally, I speculated that zip code could be an indirect estimate of family resources both in terms of median income as well as distance traveled to obtain services. Each of these factors is discussed in more depth below.

Receptive language. Receptive language was highly predictive of group membership and treatment outcome. Parent report of receptive language proved a better indicator of outcome, although this finding might be due to a higher percentage of children receiving parent report measures than clinician administered measures. Relatively higher levels of receptive language appears a *protective factor* for children with expressive language delays; 100% of children who met their goals and left treatment displayed receptive language within normal limits [OR .696, CI .531-.912]. Similarly, children who were retained in treatment were more likely to have receptive language delays of at least 25% [OR 2.684, CI 1.240-5.810], and more than half of them had receptive language delays of 40% or more.

This finding supports the literature on the importance of receptive language in children with language and developmental disabilities (cf. Camarata & Gillam, 2004). Receptive language is likely important for a number of reasons. The children in the optimal outcome group

in this study had expressive language delays equivalent to the children in other groups, so they were not simply less delayed in all areas of language. Receptive language at two years of age takes into account both following directions (i.e., pointing to objects and pictures) and pragmatics (i.e., response to name). Therefore, children with better receptive language appear more attentive and socially appropriate to parents and caregivers.

Although the transactional model of language development (Sameroff, 1979) is usually considered to be child led (i.e, the child says “ba” and the parent responds, “that’s right, a ball”), for children with language delays who may not lead the language transaction, receptive language may be crucial for parent responsivity (Gillam & Camarata, 2004) and for processing parent response “transactions.” For example, when the child is playing with a ball and the parent says, “ball,” the child who looks at the parent and points at the ball is continuing the interaction *even when the child does not vocalize*. The parent can then continue the interaction. This is a technique used in parent-mediated language intervention (Roberts & Kaiser, 2015).

In contrast, a child with poor receptive language may not respond to the parent saying “ball,” either by demonstrating joint attention (looking from the parent to the ball) or by gesturing (pointing). The language transaction is then terminated. When children do not demonstrate joint attention or vocalize, parents may have a difficult time implementing language intervention techniques. These children do make progress in treatment, but their progress is slower and the parents are less likely to see immediate gains from responsivity and interaction style.

Although receptive language is a problematic indicator of intelligence in children with receptive language delays, some of the children with low receptive language likely have cognitive impairments that are either not yet known or that are not recorded in the medical

record. Since the current study does not have a measure of nonverbal cognition, we cannot determine which children with poor receptive language might have normal intelligence, but it is likely that the children with average receptive language have average or above-average cognition.

Finally, an interesting phenomenon exists within the non-attender group with regard to receptive language. Children in the non-attender group were less likely to demonstrate receptive language delays of 40% or more [OR .826 CI .685-.996]. That is, children in the non-attender group often had receptive language delays of 25% or more, but no children in the non-attender group demonstrate a receptive language delay of 40% or more. It seems that parents of children with receptive language delays of 40% or more continued in treatment regardless of medical history or resource access.

Medical history. The ordinal variable of medical history events was not a predictive variable of eligibility, enrollment, or outcome. The nominal variable of “extensive medical history” (medical history events greater than four) was predictive only as a protective factor (“Goals Met” outcome group).

That medical history was minimally predictive in this sample does not suggest that children with complex medical needs are equally successful in speech and language treatment as children without complex medical needs. Rather, it is likely that the variable used in this study does not accurately capture medical history. Qualitative inspection of the data reveals a high number of children with a history of ear infections (n=78) and/or allergies (n=26), neither of which is associated with language outcome (Bishop & Edmundson, 1986; Paradise et al., 2001). In contrast, premature birth (n=36; Barre et al., 2011; Sansavini et al., 2010), cerebral palsy (n=5; Odding et al., 2006), and sensorineural hearing loss (n=10; Moeller, 2000; Wake et al., 2004) are

all associated with poorer language outcomes. All individual ICD codes are available in Appendix B: ICD-9 Codes Found in the Medical Charts. Before discounting medical history, it would be necessary to construct and test different versions of a medical history variable.

Access to economic resources. Median income by zip code was predictive of group membership only for those children who were successful in treatment. However, when calculating odds ratios, a nominal variable was required. The nominal variable chosen was income 1.5x the federal poverty line. When this variable was used, children in the optimal outcome group were less likely to come from low-income zip codes, while non-attenders were more likely to come from low-income zip codes. There was no interaction between income by zip code and distance from VBWC. However, the five most common zip codes with incomes less than 1.5x the federal poverty line all came from within six miles of the VBWC; 73% of non-attenders came from these zip codes.

Two additional access to resources variables were considered. Insurance status was not predictive of group membership. This is unsurprising given that half of the children in the sample used public insurance. Finally, data was obtained on whether or not children were receiving additional services (occupational therapy, physical therapy, etc). Children in the non-attender group were less likely than children in the rest of the sample to be receiving additional services. Note that this does not mean children were not eligible for additional services; after all, by the end of the study, children in the non-attender group were not receiving speech/language services, either.

The interaction between access to resources and receptive language is interesting. Recall that children in the non-attender group are as likely to have receptive language delays as the rest of the sample, but *less likely* to have receptive language delays of 40% or more. Considered

differently, children with delays of 40% or more continued in treatment (ending up in either the “remain in treatment group” or the “discharge to schools” group), regardless of access to resource variables. We might hypothesize that families with low resources sometimes choose to prioritize other concerns when their children demonstrate mild delays, but when children demonstrate severe delays, they prioritize language treatment. It is unknown if this finding is specific to toddlers (the current study includes only children 18-30 months at intake).

Economic status is often associated with minority status in studies of children (Bradley & Corwyn, 2002). Several seminal studies of language development found that children from minority race and/or ethnic groups were over-identified as having language and learning disabilities (see Artiles, Harry, Reschly, & Chinn, 2002, for a review). However, more recently, children from minority race and/or ethnic groups have been found to be under-identified and under-served (Wiley, 2013). It is unclear if these discrepancies are caused by differences in data collection or location, or due to shifting perceptions of clinical need (i.e., dialectal differences). In the current study, there were no differences in group comparisons for any variables related to race or ethnicity. However, children from minority race and/or ethnic groups were slightly more likely to be referred to treatment [OR 1.350, CI 1.064-1.712]. Importantly, they were no more or less likely to enroll in services, nor were they more likely to belong to a specific discharge group. That is, there were no disparities in outcome across racial or ethnic groups.

Length of Treatment

It is possible that some children within the “Goals Met” group might have been “false positives,” or that they were referred to treatment but in actuality did not need to be treated. If

this were so, we expected that these children would have been discharged from treatment more quickly. However, survival analysis (Figure 4. **Hazard function for “Goals Met” group.**) suggests that children in the Goals Met group were nearly all (89%) discharged after six months of treatment or longer. It is possible that some of these children were false positives but were retained in treatment for at least six months. It is also possible that the initial evaluation produced more false negatives (children who were not referred to treatment but who should have been), a “sensitivity or specificity” problem (Diepeveen et al., 2015). There were only two children in the data set who were not referred to treatment but who received follow up at Vanderbilt, both of whom later received services. This finding is suggestive of false negatives in the sample. However, the ratio of false positives to false negatives cannot be determined given the limitations of the follow-up sample.

We also expected length of treatment to vary within the “Non-Attender” group. Because only children who attended at least five sessions were considered enrolled in treatment, the non-attenders in this sample are particularly interesting. Families might choose to withdraw from treatment within the first five sessions for any number of reasons: they found another provider whom they liked better, who was available sooner, who was covered by their insurance, or who was closer to their home. Or, families may have chosen to withdraw from services due to the same reasons they withdrew from services between weeks 6-52. The “Non-Attender” group is unlikely to have withdrawn from treatment due to finding another provider (those who did through school eligibility are found in the “Discharged to Schools” group), so they could, in the future, provide insight as to the external (resources) and internal (“buy-in”) reasons that a family withdraws from services. Because of this, we expected there to be “spurts” of discharge from treatment due to non-attendance, i.e., the family attends for three months and then decides that

treatment is not useful. However, as seen in Figure 5. **Hazard function for “Non-Attender” group.**, discharge due to non-attendance occurs in a linear manner. The “Discharge to Schools” group (Figure 6), unsurprisingly, does not begin discharge until nearly 8 months after beginning treatment: the time at which these children would begin to reach 36 months and become eligible for services through the school system.

Public Health Implications

Medical chart reviews allow researchers to examine diagnosis, process, prognosis, and outcome for a community sample of children with speech and language disorders across time: children who are being treated with the current standards in practice of care. This longitudinal information has direct public health implications for adults with a history of speech and language disorders.

Longitudinal studies of language delays and disorders demonstrate that adults with a history of language disorders have poorer social and vocational outcomes than IQ- and SES-matched comparison samples (Clegg, Hollis, Mawhood & Rutter, 2004). Specifically, when compared to their own siblings, adults with histories of language disorders had long periods of unemployment (41% currently unemployed in their mid-thirties; 65% history of at least two years unemployment), often failed to live independently (59%), and had a paucity of friendships or romantic relationships. This occurred despite the fact that the adults in this study had nonverbal IQs within normal limits as adults. Similar findings are true of adults with intellectual disabilities (Hall et al., 2005) and autism spectrum disorders (Mawhood, Howlin, & Rutter, 2000; Whitehouse, Watt, Line, & Bishop, 2009).

The age of initial treatment of language disorders (comorbid or not) has decreased in recent decades; it is increasingly common for children to receive treatment before the age of three (Campbell et al., 2014; Mazurek et al., 2014). The purpose of earlier intervention is to change the trajectories, or mediate the effects of disability, for children who have speech/language disorders. While the effectiveness of early intervention programs is still unclear (cf. Camarata, 2014; Warren, 2011), all U.S. states provide some form of early intervention programming (Department of Health and Human Services, 2014). Therefore while the purpose of early intervention may not be to eliminate the effects of language disorders (which it does not appear to do; cf. Warren 2011), the purpose of early intervention is surely to improve lifelong educational, vocational, and independent living outcomes.

Limitations of the Current Study

The current study is limited by the nature of public health data and what it includes. Electronic medical records are limited as to the scope of individual goals and intervention data, and this data cannot necessarily be analyzed in a systematic way.

Electronic medical records do not typically include some of the key variables common in speech-language pathology research (e.g., an estimate of cognitive abilities). The most critical of these is nonverbal intelligence in children with speech and language disorders. Other variables that would be desirable, such as parent education are also not available in the existing medical record. It is possible for this information to be gathered in future prospective studies, although it must be gathered in ways that do not affect patient care (i.e., families must be able to opt-out). As the field of speech-language pathology moves in the direction of both retrospective and

prospective data analysis, it is likely that we will be able to collect more of the information that was not available herein.

Electronic medical records are also subject to missing data even when variables are generally available in the record. Missing data on reported variables were as high as 63% (Table 2), which was controlled for statistically. Other variables were originally collected, including performance on oral mechanism exam and participation in evaluation, yet these variables were missing so often that they could not be included for analysis. To the greatest extent possible, we handled missing data through multiple imputation or excluded variables with inconsistent data, but these are inherent limitations in chart review studies.

A second major limitation on data analysis is procedural fidelity and construct validity for the measures and methods used in assessment. Twelve clinicians were responsible for initial evaluations of the 198 children in the data set. These clinicians were all licensed speech-language pathologists with at least one year in the field and are—no doubt—highly skilled. However, there are no data on inter-rater reliability or on procedural fidelity of test administration or for clinical judgment. This is consistent with clinical practice, but not with laboratory research. Clinical judgement may also affect whether children received a clinician-administered measure (only 48% did), and clinician preference certainly affected which measures children were given. Finally, the clinicians who gave initial evaluations were unfamiliar with the children, but the clinicians who gave re-evaluations were the treating clinicians; that is, they had established relationships with the children and families, which could be a source of variation in the sample.

Some readers may be concerned about the likelihood of familywise error or the interpretation of significant results with p values between .05 and .01. As in any exploratory

study, caution should be taken in these interpretations. To address these concerns, we have provided effect sizes whenever possible. In addition, concordance of group comparisons and odds ratios increases confidence in interpretation of the results when there was concordance. For example, we may have increased confidence in the interpretation of “distance from VBWC” as a predictor of enrollment because both the interval variable in the group difference statistic and the nominal variable in the odds ratio were significant (both at $p < .05$).

A final concern is the generalizability of the data to the whole population of children with speech/language delays. These findings may be generalized to children who received speech/language evaluations around two years of age through a regional medical center and who attended treatment through an outpatient facility. They should not be generalized to older children, children who are seen only through the school system, or even potentially children from rural settings (as the average distance to VBWC was approximately 18 miles).

These limitations are very real, but they reflect the state of the field. Clinicians in schools and clinics may not have access to variables of interest and procedural fidelity will likely differ across clinicians. Moreover, causes that cannot be identified at an individual or subgroup level may sometimes be identified on a population level (Rose, 2001). Although electronic medical record analysis does not have the precision of variables that a laboratory based study can have, it allows for an overview of broad-category variables across a large number of individuals, diagnostic categories, and time.

Future Studies: Prospective Data Analysis

Despite the limitations of retrospective data analysis, the use of electronic medical records in the research of process, outcome, and eventually treatment is a relatively new arena of research in speech-language pathology. Outcome and treatment data is currently available from clinical samples (traditional treatment studies), population-based studies, national surveys, and national databases (Rosenbaum & Simon, 2016). However, the author of this paper could find very few studies of pediatric speech or language that employed medical chart reviews, although this data analysis is available in other arenas of speech-language pathology: voice disorders (i.e., Portone, Johns, & Hapner, 2008), stroke rehabilitation (i.e., Martino et al., 2005; Mauldin et al., 2005), and traumatic brain injury (i.e., Colantonio et al., 2004). History of childhood speech/language disorders is considered in other disciplines, such as Beitchman et al.,'s 2001 study of psychiatric outcomes of children with a history of speech/language impairment, a study that employed medical chart review. Medical chart review at individual institutions is particularly promising because each institution can prospectively define variables and procedures to enhance the reliability and validity of those variables.

The field of speech-language pathology has the tools to do both prospective and retrospective studies using medical chart reviews as a foundation. Prospective data analysis – that is, collecting data in clinical settings with the express intent of analyzing it systematically – allows for solutions to most, though not all, of the limitations of retrospective data analysis. Because the variables are determined in advance, clinicians can make a greater effort to collect specific variables, helping both to include necessary variables in the analysis and to decrease the extent of missing data within those variables. Procedural fidelity can be tested intermittently, and clinicians are made aware that their procedures both affect individual children and the

overall clinical outcomes of the sample. Sets of measures can be chosen as an ideal diagnostic battery and deviated from as appropriate, according to clinical judgement, for better direct comparison.

Ideally, we could create a set of variables for every child who entered the clinic. Although the current study found significant results using median income by zip code, parent reports of income and educational status would be better indicators of individual family resources. Adding these variables requires only alteration of a few questions in the intake packet that families already fill out during initial evaluation. It seems clear that children would benefit from receiving both a parent-reported measure and a clinician-administered measure (preferably not the PLS-5); information on overall developmental ability (a proxy for cognition in young children) could be made available using scales such as the Mullen, Reynell, or Bayley, among others. Information on social reciprocity and joint attention can easily be collected using parent reports such as the Childhood Autism Rating Scale (CARS) or through an elicited sample such as the Autism Diagnostic Observation Schedule (ADOS-2).

Figure 7. **Prospective data analysis theoretical model.** depicts a theoretical model of language outcome. Some variables, such as initial language abilities, medical history, and resource factors, were available in the current study, although these variables could likely be improved. Other variables, such as cognition, could become available at the initial evaluation, should the institution in question chose to pursue these variables. These variables can all be considered initial predictor variables.

More interesting are evaluations of different treatment strategies (and ensuing response-to-treatment). The current study was not designed to test to effectiveness of treatment strategies, nor were the available data in the medical records to allow for testing these questions.

Nonetheless, these are prime effectiveness questions: what is the relationship of a given treatment strategy (or set of strategies) on outcome given a set of initial factors? Who benefits from a given treatment strategy based on their initial characteristics? There are multiple ways that treatment strategies can be defined, both through observational measurement and through annotations in the medical record. Moreover, the length of treatment in a clinical setting allows for strategies to change across time, as the child learns necessary skills to move forward with communication. Prospective effectiveness studies would be particularly useful to address these questions.

Finally, more information about attendance and therapeutic alliances between families and clinicians is needed. It is not enough to simply look at non-attenders and speculate as to why they might fail to come to treatment sessions. Accurate attendance information that links the specific reasons for poor attendance is needed. From there, we need to identify which families are at risk of failure to attend and if there is anything that we, as clinicians, can do to improve the quality of care for these families. Although counseling and therapeutic relationship are discussed within speech-language-pathology graduate programs, there are few indicators of therapeutic relationships that can be used within clinical settings (cf. Plexico, Manning, & DiLollo, 2010).

Conclusion

There is a pressing need for effectiveness studies of “real world” community based clinical outcomes. This is well understood—and studied—in medicine, but rarely examined in allied health fields such as pediatric speech language pathology. This study, of one clinic in one city, generated 198 cases, which makes it a large outcome study for speech-language pathology.

If information on treatment strategies were available, it could have been one of the largest “intervention” studies in the literature. There is no doubt that tremendous data mining opportunities exist in clinics across the country, especially as electronic medical records become mandatory—and more widely available.

Despite the inherent obstacles to conducting the research and limitations in the variables, interpretable and hypothesis driven analyses were conducted—and yielded credible results for the parameters that could be explored (e.g., receptive language level). It is clear that studies of this nature can advance the knowledge base and the evidence base of the field and form an important complement to efficacy studies.

REFERENCES

- 2013 Federal Poverty Guidelines (2016, March 22). Department of Health and Human Services. Retrieved from: <https://aspe.hhs.gov/2013-poverty-guidelines>.
- 2013 Per Capita Income by Zip Code (2015, December 12). Retrieved from: <http://www.census.gov/data.html>.
- Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of experimental child psychology*, 87(2), 85-106
- Artiles, A. J., Harry, B., Reschly, D. J., & Chinn, P. C. (2002). Over-identification of students of color in special education: A critical overview. *Multicultural Perspectives*, 4(1), 3-10.
- American Speech-Language-Hearing Association. (2007). Executive Summary for JCIH Year 2007 Position Statement: Principles and Guidelines for Early Hearing Detection and Intervention Programs. Available from www.asha.org.
- Baker, E., & McLeod, S. (2011). Evidence-based practice for children with speech sound disorders: Part 1 narrative review. *Language, Speech, and Hearing Services in Schools*, 42(2), 102-139.
- Barre, N., Morgan, A., Doyle, L.W., & Anderson, P.J. (2011). Language abilities in children who were very preterm and/or very low birth weight: A meta-analysis. *The Journal of Pediatrics*, 158(5), 766-774.
- Bedore, L. M., & Pena, E. D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, 11(1), 1-29.
- Beitchman, J.H., Wilson, B., Johnson, C.J., Atkinson, L., Young, A., Adlaf, E., Escobar, M., and Douglas, L. (2001). Fourteen-year follow-up of speech/language-impaired and control children: Psychiatric outcome. *Journal of the American Academy of Child & Adolescent Psychiatry* 40(1), 75-82.
- Bishop, D. V., & Edmundson, A. (1986). Is otitis media a major cause of specific developmental language disorders? *International Journal of Language & Communication Disorders*, 21(3), 321-338.
- Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, 53(1), 371-399.
- Briesacher, B. A., Gurwitz, J. H., & Soumerai, S. B. (2007). Patients at-risk for cost-related medication nonadherence: A review of the literature. *Journal of General Internal Medicine*, 22(6), 864-871.

- Camarata, S. (2014). Early identification and early intervention in autism spectrum disorders: Accurate and effective? *International Journal of Speech-Language Pathology*, 16(1), 1-10.
- Camarata, S. M., Nelson, K. E., & Camarata, M. N. (1994). Comparison of conversational-recasting and imitative procedures for training grammatical structures in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 37(6), 1414-1423.
- Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., & Pan, Y. (2014). Early childhood investments substantially boost adult health. *Science*, 343(6178), 1478-1485.
- Carroll, J. B. (1961). The nature of the data, or how to choose a correlation coefficient. *Psychometrika*, 26(4), 347-372.
- Cicchetti, D.V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290.
- Chapman, R. (2000). Children's language learning: An interactionist perspective. *Journal of Child Psychology and Psychiatry*, 41, 33-54.
- Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental language disorders-a follow-up in later adult life: Cognitive, language and psychosocial outcomes. *Journal of Child Psychology and Psychiatry*, 46(2), 128-149.
- Colantonio, A., Ratcliff, G., Chase, S., Kelsey, S., Escobar, M., & Vernich, L. (2004). Long term outcomes after moderate to severe traumatic brain injury. *Disability and Rehabilitation*, 26(5), 253-261.
- Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science*, 4(1), 50.
- Davis, T. N., Lancaster, H. S., & Camarata, S. (2015). Expressive and receptive vocabulary learning in children with diverse disability typologies. *International Journal of Developmental Disabilities*, 2047387715Y-0000000010.
- De Giacomo, A., & Fombonne, E. (1998). Parental recognition of developmental abnormalities in autism. *European Child & Adolescent Psychiatry*, 7(3), 131-136.
- Diepeveen, F. B., Dusseldorp, E., Bol, G. W., Oudesluys-Murphy, A. M., & Verkerk, P. H. (2015). Failure to meet language milestones at two years of age is predictive of specific language impairment. *Acta Paediatrica*.
- Durning, A. T. (1992). *How much is enough?: the consumer society and the future of the earth*. WW Norton & Company.

- Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American Statistical Association*, 83(402), 414-425.
- Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, 83(1), 203-222.
- Fischer, M. A., Choudhry, N. K., Brill, G., Avorn, J., Schneeweiss, S., Hutchins, D., . . . Shrank, W. H. (2011). Trouble getting started: Predictors of primary medication nonadherence. *The American Journal of Medicine*, 124(11), 1081-1089.
- Geronimus, A.T., Bound, J. (1998). Use of census-based aggregate variables to proxy for socioeconomic group: Evidence from national samples. *American Journal of Epidemiology*, 148(5), 475-486.
- Gillum, H., & Camarata, S. (2004). Importance of treatment efficacy research on language comprehension in MR/DD research. *Mental Retardation and Developmental Disabilities Research Reviews*, 10(3), 201-207.
- Hall, I., Strydom, A., Richards, M., Hardy, R., Bernal, J., & Wadsworth, M. (2005). Social outcomes in adulthood of children with intellectual impairment: Evidence from a birth cohort. *Journal of Intellectual Disability Research*, 49(3), 171-182.
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23–34.
- Hartmann, D. P. (1977). Considerations in the choice of interobserver reliability estimates. *Journal of Applied Behavior Analysis*, 10(1), 103-116.
- Harris, P.A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., Conde, J.C. (2009). Research electronic data capture (REDCap) – A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, 42(2), 377-381.
- Howlin, P., Magiati, I., & Charman, T. (2009). Systematic review of early intensive behavioral interventions for children with autism. *American journal on intellectual and developmental disabilities*, 114(1), 23-41.
- Howlin, P., Mawhood, L., & Rutter, M. (2000). Autism and developmental receptive language disorder—A follow-up comparison in early adult life. II: Social, behavioural, and psychiatric outcomes. *Journal of Child Psychology and Psychiatry*, 41(05), 561-578.
- IDEA 2004: Building the Legacy (2015, October 4). Retrieved from: <http://idea.ed.gov/part-c/search/new>.
- Justice, L. M., Meier, J., & Walpole, S. (2005). Learning New Words From Storybooks: An Efficacy Study With At-Risk Kindergartners. *Language, speech, and hearing services in schools*, 36(1), 17-32.

- Kasari, C., Freeman, S., & Paparella, T. (2006). Joint attention and symbolic play in young children with autism: A randomized controlled intervention study. *Journal of Child Psychology and Psychiatry*, *47*(6), 611-620.
- Katz, G., & Lazcano-Ponce, E. (2008). Intellectual disability: Definition, etiological factors, classification, diagnosis, treatment and prognosis. *Salud Pública de México*, *50*, 132-141.
- Kentner, J., & Miller, N. (2009). Effects of visible vs concealed alphabet cues on speech intelligibility in dysarthria. *International Journal of Therapy & Rehabilitation*, *16*(5).
- Klein, J. P., & Moeschberger, M. L. (2005). *Survival analysis: techniques for censored and truncated data*. Springer Science & Business Media.
- Kolstoe, O. P. (1958). Language training of low-grade mongoloid children. *American Journal of Mental Deficiency*, *63*(1), 17-30.
- Law, J., Garrett, Z., & Nye, C. (2004). The efficacy of treatment for children with developmental speech and language delay/disorders meta-analysis. *Journal of Speech, Language, and Hearing Research*, *47*(4), 924-943.
- Lehane, E., & McCarthy, G. (2007). Intentional and unintentional medication non-adherence: a comprehensive framework for clinical research and practice? A discussion paper. *International Journal of Nursing Studies*, *44*(8), 1468-1477.
- Magnuson, K. A., & Duncan, G. J. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps?. *The Future of Children*, *15*(1), 35-54.
- Martino, R., Foley, M., Bhogal, S., Diamant, N., Speechley, M., and Teasell, R. (2005). Dysphagia after stroke incidence, diagnosis, and pulmonary complications. *Stroke* *36*, 2756-2763.
- Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning a meta-analysis. *Review of Educational Research*, *80*(3), 300-335.
- Maulden, S. A., Gassaway, J., Horn, S. D., Smout, R. J., & DeJong, G. (2005). Timing of initiation of rehabilitation after stroke. *Archives of Physical Medicine and Rehabilitation*, *86*(12), 34-40.
- Mawhood, L., Howlin, P., & Rutter, M. (2000). Autism and developmental receptive language disorder—A comparative follow-up in early adult life. I: Cognitive and language outcomes. *Journal of Child Psychology and Psychiatry*, *41*(5), 547-559.
- Mazurek, M. O., Handen, B. L., Wodka, E. L., Nowinski, L., Butter, E., & Engelhardt, C. R. (2014). Age at first autism spectrum disorder diagnosis: the role of birth cohort, demographic factors, and clinical features. *Journal of Developmental & Behavioral Pediatrics*, *35*(9), 561-569.

- Meyer, C., Witte, J., Hildmann, A., Hennecke, K. H., Schunck, K. U., Maul, K., ... & Hartmann, S. (1999). Neonatal screening for hearing disorders in infants at risk: incidence, risk factors, and follow-up. *Pediatrics*, *104*(4), 900-904.
- Miller, J. (1981). *Assessing Language Production in Children*. Boston, MA: Allyn & Bacon.
- Moeller, M. P. (2000). Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics*, *106*(3).
- Næss, K. A. B., Lyster, S. A. H., Hulme, C., & Melby-Lervåg, M. (2011). Language and verbal short-term memory skills in children with Down syndrome: A meta-analytic review. *Research in Developmental Disabilities*, *32*(6), 2225-2234.
- Odding, E., Roebroek, M. E., & Stam, H. J. (2006). The epidemiology of cerebral palsy: Incidence, impairments and risk factors. *Disability and Rehabilitation*, *28*(4), 183-191.
- Olswang, L., Bain, B. (2013). Treatment research. Golper, L. A., & Fratalli, C. (Eds.). *Outcomes In Speech Language Pathology*, *2*, 245-264.
- Olswang, L. B., & Prelock, P. A. (2016). Bridging the gap between research and practice: Implementation science. *Journal of Speech, Language, and Hearing Research*, *58*(6), 1818-1826.
- Paradise, J. L., Feldman, H. M., Campbell, T. F., Dollaghan, C. A., Colborn, D. K., Bernard, B. S., . . . Kurs-Lasky, M. (2001). Effect of early or delayed insertion of tympanostomy tubes for persistent otitis media on developmental outcomes at the age of three years. *New England Journal of Medicine*, *344*(16), 1179-1187.
- Plexico, L.W., Manning, W.H., and DiLollo, A. (2010) Client perceptions of effective and ineffective therapeutic alliances during treatment for stuttering. *Journal of Fluency Disorders* *35*(4), 333-354.
- Portone, C., Johns, M. M., & Hapner, E. R. (2008). A review of patient adherence to the recommendation for voice therapy. *Journal of Voice*, *22*(2), 192-196.
- Roberts, J. E., Mankowski, J. B., Sideris, J., Goldman, B. D., Hatton, D. D., Mirrett, P. L., . . . Bailey, D. B. (2008). Trajectories and predictors of the development of very young boys with fragile X syndrome. *Journal of Pediatric Psychology*.
- Roberts, M. Y., & Kaiser, A. P. (2015). Early intervention for toddlers with language delays: A randomized controlled trial. *Pediatrics*, *135*(4), 686-693.
- Rose, G. (2001). Sick individuals and sick populations." *International Journal of Epidemiology* *30*(3), 427-432.
- Rosenbaum, S. & Simon, P. (2016). *Speech and Language Disorders in Children: Implications for the Social Security Administration's Supplemental Security Income Program*. National Academy Press: Washington, DC.

- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (231-244). New York, NY: Sage.
- Sameroff, A. (1979, March). Theoretical and empirical issues in the operationalization of transactional research. In *Meeting of the Society for Research in Child Development*. San Francisco.
- Sansavini, A., Guarini, A., Justice, L.M., Savini, S., Broccoli, S., Alessandrini, R., & Faldella, G. (2010). Does preterm birth increase a child's risk for language impairment? *Early Human Development*, 86(12), 765-772.
- Shavers, V. L. (2007). Measurement of socioeconomic status in health disparities research. *Journal of the National Medical Association*, 99(9), 1013.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420.
- Šidák, Z. (1967). Rectangular confidence region for the means of multivariate normal distributions. *Journal of the American Statistical Association*, 62, 626–633.
- Spreckley, M., & Boyd, R. (2009). Efficacy of applied behavioral intervention in preschool children with autism for improving cognitive, language, and adaptive behavior: A systematic review and meta-analysis. *The Journal of Pediatrics*, 154(3), 338-344.
- Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., ... & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *British Journal of Medicine*, 338, b2393.
- Strauss, K., MacLean, C., Troy, A., & Littenberg, B. (2006). Driving distance as a barrier to glycemic control in diabetes. *Journal of General Internal Medicine*, 21(4), 378-380.
- Sutera, S., Pandey, J., Esser, E. L., Rosenthal, M. A., Wilson, L. B., Barton, M., ... & Fein, D. (2007). Predictors of optimal outcome in toddlers diagnosed with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 37(1), 98-107.
- Swisher, L., Plante, E., & Lowell, S. (1994). Nonlinguistic deficits of children with language disorders complicate the interpretation of their nonverbal IQ scores. *Language, Speech, and Hearing Services in Schools*, 25(4), 235-240.
- Tabachnick B.G., & Fidell, L.S. (2001). *Using Multivariate Statistics*. Needham Heights, Mass, Allyn & Bacon.
- Taylor, L. E., Swerdfeger, A. L., & Eslick, G. D. (2014). Vaccines are not associated with autism: an evidence-based meta-analysis of case-control and cohort studies. *Vaccine*, 32(29), 3623-3629.
- TennCare Kids (2015, October 4). Retrieved from:
<https://www.tn.gov/tenncare/section/tenncare-kids>.

- Thurm, A., Lord, C., Lee, L. C., & Newschaffer, C. (2007). Predictors of language acquisition in preschool children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 37(9), 1721-1734.
- Tomblin, J. Bruce, Nancy L. Records, Paula Buckwalter, Xuyang Zhang, Elaine Smith, and Marlea O'Brien. "Prevalence of specific language impairment in kindergarten children." *Journal of Speech, Language, and Hearing Research* 40, no. 6 (1997): 1245-1260.
- van der Schuit, M., Segers, E., van Balkom, H., & Verhoeven, L. (2011). How cognitive factors affect language development in children with intellectual disabilities. *Research in Developmental Disabilities*, 32(5), 1884-1894.
- Wake, M., Tobin, S., Cone-Wesson, B., Dahl, H. H., Gillam, L., McCormick, L., . . . Williams, J. (2006). Slight/mild sensorineural hearing loss in children. *Pediatrics*, 118(5), 1842-1851.
- Watson, L. R., Crais, E. R., Baranek, G. T., Dykstra, J. R., & Wilson, K. P. (2013). Communicative gesture use in infants with and without autism: A retrospective home video study. *American Journal of Speech-Language Pathology*, 22(1), 25-39.
- Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-VanderWeele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics*, 127(5), 1303-1311.
- Whitehouse, A. J., Watt, H. J., Line, E. A., & Bishop, D. V. (2009). Adult psychosocial outcomes of children with specific language impairment, pragmatic language impairment and autism. *International Journal of Language & Communication Disorders*, 44(4), 511-528.
- Wild, K. T., Betancourt, L. M., Brodsky, N. L., & Hurt, H. (2013). The effect of socioeconomic status on the language outcome of preterm infants at toddler age. *Early Human Development*, 89(9), 743-746.
- Wiley, T. G. (2013). A brief history and assessment of language rights in the United States. *Language Policies in Education: Critical Issues*, 61-90.
- Yoder, P., Woynaroski, T., Fey, M., & Warren, S. (2014). Effects of dose frequency of early communication intervention in young children with and without Down syndrome. *American Journal on Intellectual and Developmental Disabilities*, 119(1), 17-32.

TABLES

Table 1. Specific aims, hypotheses, statistical procedures, and degrees of freedom.

Specific Aim	Hypotheses	Statistical Procedure	Degrees of Freedom Used
Aim 1	Children who are evaluated and judged not to need services will (a) perform better on standard measures and (b) demonstrate fewer medical concerns than children who are evaluated and judged to need services.	t-test or Mann-Whitney U; Cohen's d (effect size)	2
	Children who (a) receive low language scores and (b) have medical concerns will be more likely to be referred to treatment.	Odds ratios	2
Aim 2	Children who enroll in treatment at Vanderbilt will (a) live closer to Vanderbilt as measured by distance by zip code and (b) have greater medical concerns than children who do not enroll in treatment at Vanderbilt.	t-test or Mann-Whitney U; Cohen's d (effect size)	2
	Children who (a) live within 20 miles of VBWC and (b) have medical concerns will be more likely to enroll in treatment.	Odds ratios	2
Aim 3	Children who are discharged from treatment due to meeting goals or reaching normal limits will differ from children who remain in treatment based on having (a) higher receptive language, (b) fewer medical concerns, and (c) higher income by zip code.	ANOVA	3
	Children who (a) have receptive language within normal limits, (b) have no medical concerns, and (c) come from zip codes with incomes above 1.5 times the federal poverty line will be more likely to be discharged due to meeting goals or reaching normal limits.	Odds ratio	3

Table 2. Reliability and missing data.

Measure	Reliability (Percent Agreement)	Percent Missing Data	Imputed	Effect of Imputation on Results
Demographic and Medical Variables				
Age	100%	0%		
Race	100%	22%	X	No effect
Ethnicity	95%	20%	X	No effect
Primary Language	97%	0%		
Medical History Index	ICC=.701	2%		
Family History Index	ICC=.614	2%		
Median Income By Zip Code	100%	0%		
Insurance Status	100%	0%		
Distance from Vanderbilt	92%	0%		
Time to Discharge	95%	1%		
Speech/Language Variables				
CA ^a Receptive Language	100%	52% ^c	X	No effect
CA Expressive Language	100%	52% ^c	X	Affected enrollment (p>.05 on raw data and p<.05 on imputed data)
PR ^b Expressive Language	100%	37% ^c	X	No effect
PR Expressive Language	100%	37% ^c	X	No effect
Number of Different Words	85%	50%	X	No effect
Percent Intelligibility	85%	68%	X	No effect

^aCA = clinician administered. ^bPR = parent report (standardized). ^cThe proportion of missing data for these variables may seem high, but this is caused by children receiving either a parent-reported measure or a clinician-administered measure but not both. In fact, only 12% of children received both a parent-reported measure and a

clinician-administered measure. 90% of children received standardized scores on at least one measure. Therefore the construct “standardized scores” has only 10% missing data. Nevertheless, missing data for each individual variable is presented in this table.

Table 3. Variable types by time of administration and level of measurement.

Measure	Administration			Variable Type	
	Administered at Intake	Administered at Outcome	Derived	Nominal/ Ordinal	Interval
Demographic and Medical Variables					
Age	X	X			X
Race	X			X	
Ethnicity				X	
Primary Language	X			X	
Medical History Index			X	X	
Family History Index			X	X	
Median Income By Zip Code	X		X		X
Insurance Status	X			X	
Distance from Vanderbilt			X	X	X
Speech/Language Variables					
Clinician-Administered Language Measures	X	X		X	X
Parent-Reported Language Measures	X	X		X	X
Clinician Administered Articulation Measures		X			X
Number of Different Words	X				X
Percent Intelligibility	X				X
Status Variables					
Referred to Treatment	X			X	
Treated at Vanderbilt	X			X	
Discharge Status		X		X	
Days to Discharge			X		X

Table 4. Kurtosis and skew for intake language variables.

	Whole Sample		Treatment Group	
	Skew	Kurtosis	Skew	Kurtosis
Clinician-Administered Receptive SS	-.008	-.719	.031	-.857
Clinician-Administered Expressive SS	.319	.584	-.267	.558
Parent Reported Receptive SS	.144	-.388	.247	-.382
Parent Reported Expressive SS	.236	-1.045	.314	-1.130

Table 5. Referral to treatment based on initial evaluation variables.

	Children Referred for Treatment (n=154)	Children Not Referred for Treatment (n=39)	Statistical Values
Demographic Variables			
Age (months)	24.70 (3.52)	24.59 (4.33)	n.s.
Race	49% minority	39% minority	n.s.
Gender	75.3% male	71.8% male	n.s.
Median Income by Zip Code	\$42,104 (13,077)	\$43,746 (13,642)	n.s.
Distance from Vanderbilt by Zip Code (miles)	24.73 (25.11)	26.71 (27.01)	n.s.
Public Insurance (TennCare)	48%	44%	n.s.
Receiving Additional Services (PT, OT, TEIS)	39%	31%	n.s.
Speech/Language Variables			
Clinician-Administered Receptive SS	77.80 (15.34)	101.44 (11.40)	t=3.14, p<.01, d=1.95
Clinician-Administered Expressive SS	78.60 (10.06)	101.19 (8.67)	t=5.29, p<.01, d=2.41
Parent Report Receptive SS	79.36 (16.96)	94.58 (14.31)	t=2.12, p<.05, d=.97
Parent Report Expressive SS	68.22 (11.35)	85.52 (8.99)	t=6.26, p<.01, d=1.69
Parent Reported Number of Different Words	7.20 (7.94)	22.63 (20.36) n=16	t=5.21, p<.01, d=1.00
Parent Reported Percent Intelligibility	24.72% (30.00)	33.00% (34.66)	n.s.
Medical Variables			
Medical History Index	1.84 (2.06)	1.54 (1.50)	n.s.
Family History Index	0.36 (0.48)	0.47 (0.65)	n.s.

Table 6. Odds of referral to treatment for each nominal variable.

	Odds Ratio	Lower CI	Upper CI
Referral to Treatment			
Race	1.350*	1.064	1.712
Gender	.953	.873	1.041
Income Below 1.5x Poverty Line	.817	.588	1.135
Distance from Vanderbilt >20 miles	1.048	.854	1.285
Private Insurance	.913	.781	1.068
Patient Medical Hx	.807	.559	1.097
History of Otitis Media (n=78)	.721*	.582	.893
History of Premature Birth (n=36)	1.393*	1.006	1.927
Family History	.929*	.773	.994
Additional Services	.827	.673	1.016

*significant variable based on χ^2 ; confidence interval does not cross 1.0

Table 7. Enrollment in treatment for 154 children who were made eligible based on initial evaluation variables.

	Children Treated at Vanderbilt (n=67)	Children Lost to Follow Up (n=87)	Statistical Values
Demographic Variables			
Age (months)	24.40 (3.68)	24.67 (3.66)	n.s.
Race	38% minority	32% minority	n.s.
Gender	75% male	75% male	n.s.
Median Income by Zip Code	\$42,620 (14,445)	\$42,358 (12,483)	n.s.
Distance from Vanderbilt by Zip Code (miles)	18.84 (18.89)	28.47 (27.84)	t=2.54, p<.05, d=0.43
Public Insurance (TennCare)	49%	51%	n.s.
Receiving Additional Services (PT, OT, TEIS)	45%	34%	n.s.
Speech/Language Variables			
Clinician Administered Receptive SS	80.43 (15.44)	84.85 (18.06)	n.s.
Clinician Administered Expressive SS	78.32 (10.42)	86.66 (14.04)	t=2.64, p<.05, d=.67
Parent Reported Receptive SS	79.98 (17.80)	83.21 (17.18)	n.s.
Parent Reported Expressive SS	67.23 (11.35)	73.36 (12.85)	n.s.
Parent Reported Number of Different Words	7.54 (8.12)	10.61 (13.14)	n.s.
Parent Reported Percent Intelligibility	22.74% (26.91)	26.70% (31.14)	n.s.
Medical Variables			
Medical History Index	1.76 (1.98)	1.77 (1.95)	n.s.
Family History Index	0.54 (0.92)	0.54 (0.80)	n.s.

Table 8. Odds of enrollment in treatment for each nominal variable.

	Odds Ratio	Lower CI	Upper CI
Enrollment in Treatment			
Race	1.094	.999	1.198
Gender	1.000	.932	1.073
Income Below 1.5x Poverty Line	1.316*	1.024	1.692
Distance from Vanderbilt >20 miles	.549*	.444	.577
Private Insurance	.961	.851	1.085
Patient Medical Hx	.958	.716	1.282
History of Otitis	.559*	.437	.715
Media (n=45)			
History of	1.018	.796	1.305
Premature Birth (n=23)			
Family History	.880	.661	1.343
Additional Services	1.342*	1.159	1.555

*significant variable based on χ^2 ; confidence interval does not cross 1.0

Table 9. Discharge based on initial and follow up evaluation variables.

	Children Discharged Within 1 Year: Goals Met (n=9)	Children Discharged Within 1 Year: School Setting (n=14)	Children Who Remain in Treatment At 1 Year (n=26)	Children Discharged Within 1 Year: Non-Attenders (n=11)
Demographic Variables				
Age (months)	25.88 (3.09)	24.07 (4.34)	23.35 (4.35)	24.80 (3.16)
Race	25% minority	50% minority	50% minority	40% minority
Gender	55% male	50% male	77% male	91% male
Median Income by Zip Code	\$49,622 (19,067)*	\$36,833 (8,309)	\$43,741 (16,086)	\$36,308 (8,297)
Distance from Vanderbilt by Zip Code (miles)	16.79 (10.07)	18.76 (27.25)	21.16 (21.35)	12.56 (6.90)
Public Insurance (TennCare)	50%	64%	54%	70%
Receiving Additional Services (PT, OT, TEIS)	48%	50%	54%	9%*
Speech/Language Variables				
Intake Clinician Administered Receptive SS	85.50 (13.24)	74.57 (11.98)	71.86 (10.96)	82.75 (20.32)
Intake Clinician Administered Expressive SS	79.50 (3.54)	75.29 (10.48)	79.25 (12.37)	81.25 (11.59)
Intake Parent Reported Receptive SS	94.83 (8.61)*	79.83 (10.13)	70.82 (15.01)*	85.00 (12.00)
Intake Parent Reported Expressive SS	71.17 (6.88)	70.83 (10.63)	63.88 (12.35)*	71.50 (14.82)
Parent Reported Number of Different Words	13.00 (16.97)	6.00 (7.13)	4.00 (4.88)*	12.00 (5.70)
Parent Reported Percent Intelligibility	25.00 (35.36)	--- ^a	22.78 (35.28)	25.00 (7.07)

Medical Variables				
Medical History Index	1.50 (0.76)	2.38 (1.85)	1.92 (1.95)	0.90 (1.29)
Family History Index	0.50 (0.76)	0.71 (1.13)	0.44 (0.82)	1.00 (1.50)*
Outcome Variables				
Outcome Clinician Administered Receptive SS ^b	103.56 (13.36)**	75.50 (12.76)	76.46 (21.66)	--- ^d
Outcome Clinician Administered Expressive SS ^b	98.57 (12.91)**	80.88 (7.55)	78.23 (19.41)	--- ^d
Outcome Articulation SS	89.83 (15.68)	82.67 (15.57)	--- ^c	--- ^d

^a Only three children in this category had reported percent intelligibility. Therefore, means and standard deviations are not reported. These children are included in the whole sample for the purpose of comparison to other groups. ^b Only two children received parent-reported measures at outcome (both children were in the “remain in treatment” group). Therefore, only clinician-administered measures are reported at outcome. ^c Only three children in this category received articulation measures. Therefore, means and standard deviations are not reported. ^d Non-attenders, by definition of failing to continue in clinician-recommended treatment, do not have outcome scores.

*p<.05

**p<.01

Table 10. Odds of each nominal variable within discharge group.

	Odds Ratio	Lower CI	Upper CI
Discharge Due to Goals Met			
Age	1.032	.296	3.593
Race	1.680	.390	7.237
Gender	1.458	.430	4.945
Income Below 1.5x Poverty Line	.719*	.579	.893
Private Insurance	1.250	.290	5.385
Patient Medical Hx	.775*	.656	.916
Patient Family Hx	.732	.468	1.891
Additional Services	.406	.094	1.745
25% Receptive Delay CA	.800	.587	1.091
40% Receptive Delay CA	.857	.692	1.062
25% Receptive Delay PR	.632*	.448	.890
40% Receptive Delay PR	.696*	.531	.912
Discharge To Schools			
Age	1.720	.550	5.382
Race	.480	.201	1.148
Gender	.503	.253	1.003
Income Below 1.5x Poverty Line	1.563	.502	4.862
Private Insurance	.652	.250	1.703
Patient Medical Hx	.583	.212	1.601
Patient Family Hx	.844	.322	1.549
Additional Services	1.582	.486	3.049
25% Receptive Delay CA	.632	.129	3.088
40% Receptive Delay CA	1.745	.671	4.540

25% Receptive Delay PR	.632	.129	3.088
40% Receptive Delay PR	.696	.102	4.730

Remain in Treatment

Age	.525	.275	1.003
Race	.903	.465	1.754
Gender	.902	.402	2.026
Income Below 1.5x Poverty Line	1.206	.561	2.595
Private Insurance	.755	.386	1.479
Patient Medical Hx	.882	.379	2.054
Patient Family Hx	1.179	.873	1.655
Additional Services	1.466	.752	2.857
25% Receptive Delay CA	.360	.091	1.417
40% Receptive Delay CA	1.125	.201	6.288
25% Receptive Delay PR	2.684*	1.240	5.810
40% Receptive Delay PR	3.462*	1.976	12.274

Discharge Due To Non-Attendance

Age	1.204	.359	4.041
Race	.560	.165	1.900
Gender	.325	.046	2.299
Income Below 1.5 Poverty Line	1.458*	1.045	2.250
Insurance Status	.447	.131	1.523
Patient Medical Hx	1.575	.235	10.568
Patient Family Hx	1.322*	1.074	2.135
Additional Services	.743*	.569	.970
25% Receptive Delay CA	1.800	.238	13.627

40% Receptive Delay CA	.429	.077	2.371
25% Receptive Delay PR	1.263	.154	10.388
40% Receptive Delay PR	.826*	.685	.996

*significant variable based on χ^2 ; confidence interval does not cross 1.0

Table 11. Outcomes odds ratios for receptive language.

	Odds Ratio	Lower CI	Upper CI
Discharge Due to Goals Met			
25% Receptive Delay CA	.385*	.193	.765
40% Receptive Delay CA	.500*	.306	.816
Discharge To Schools			
25% Receptive Delay CA	6.494*	1.818	23.256
40% Receptive Delay CA	3.195*	1.548	6.623
Remain in Treatment			
25% Receptive Delay CA	1.701	.748	3.861
40% Receptive Delay CA	4.900*	1.372	30.764

*significant variable based on χ^2 ; confidence interval does not cross 1.0

Note. Non-attenders, by virtue of failing to continue treatment, do not have outcome measures and are therefore not included in this table.

Table 12. Concordance of clinician-administered measures and parent-reported measures.

	Children with Clinician-Administered Assessments Only (n=54)	Children with Parent Reported Assessments Only (n=103)	Children with Both Clinician and Parent Assessments (n=21)
CA Receptive Standard Score	83.70 (17.74)	--	82.64 (16.09)
CA Expressive Standard Score	85.68 (14.12)*	--	78.44 (9.59)* **
PR Receptive Standard Score	--	81.68 (17.51)	83.92 (16.39)
PR Expressive Standard Score	--	71.53 (12.46)	66.50 13.96)**

*p<.01 horizontal comparison

**p<.01 vertical comparison

FIGURES

Figure 1. Initial plan to structure flow of participants from initial speech language evaluation to outcome. Children lost to follow up at various positions are shown to the right. Two anticipated groups at outcome.

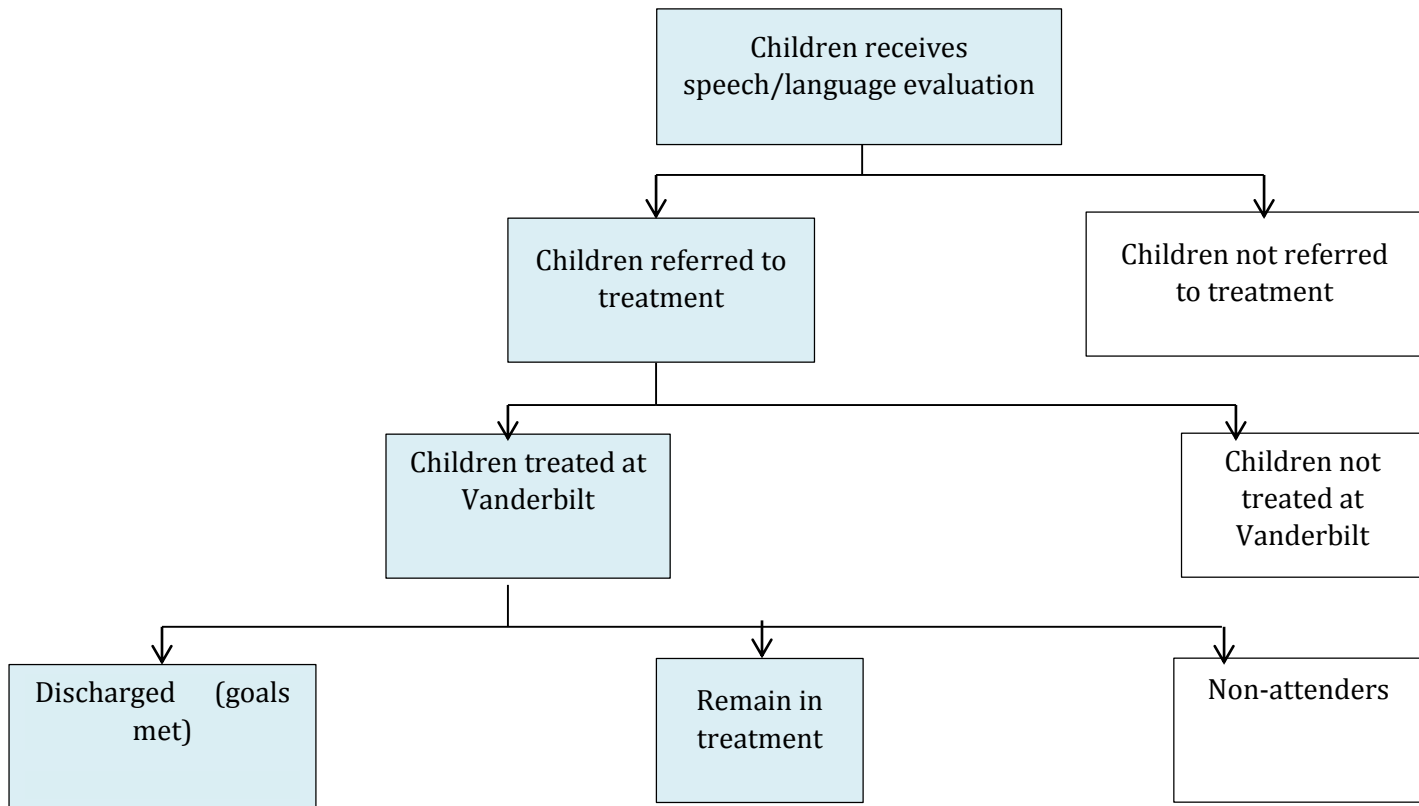


Figure 2. Final flowchart for children who received speech/language evaluations, including number of children at each stage. The original outcome groups are highlighted in darker blue.

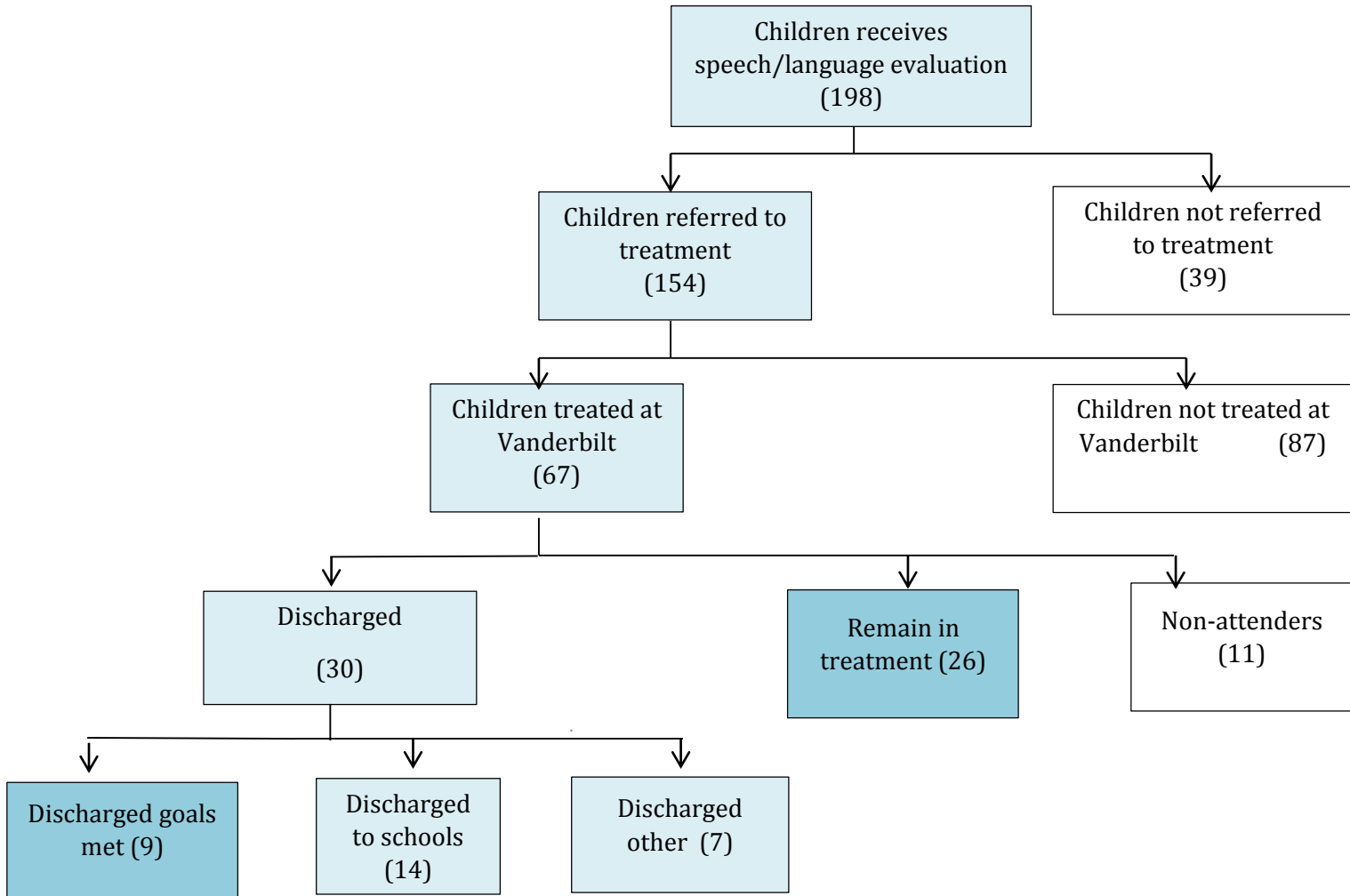


Figure 3. Determination of eligibility for the study. Records from 532 children were initially gathered. 198 children were used in the current study. An additional 58 children are included in the Appendix.

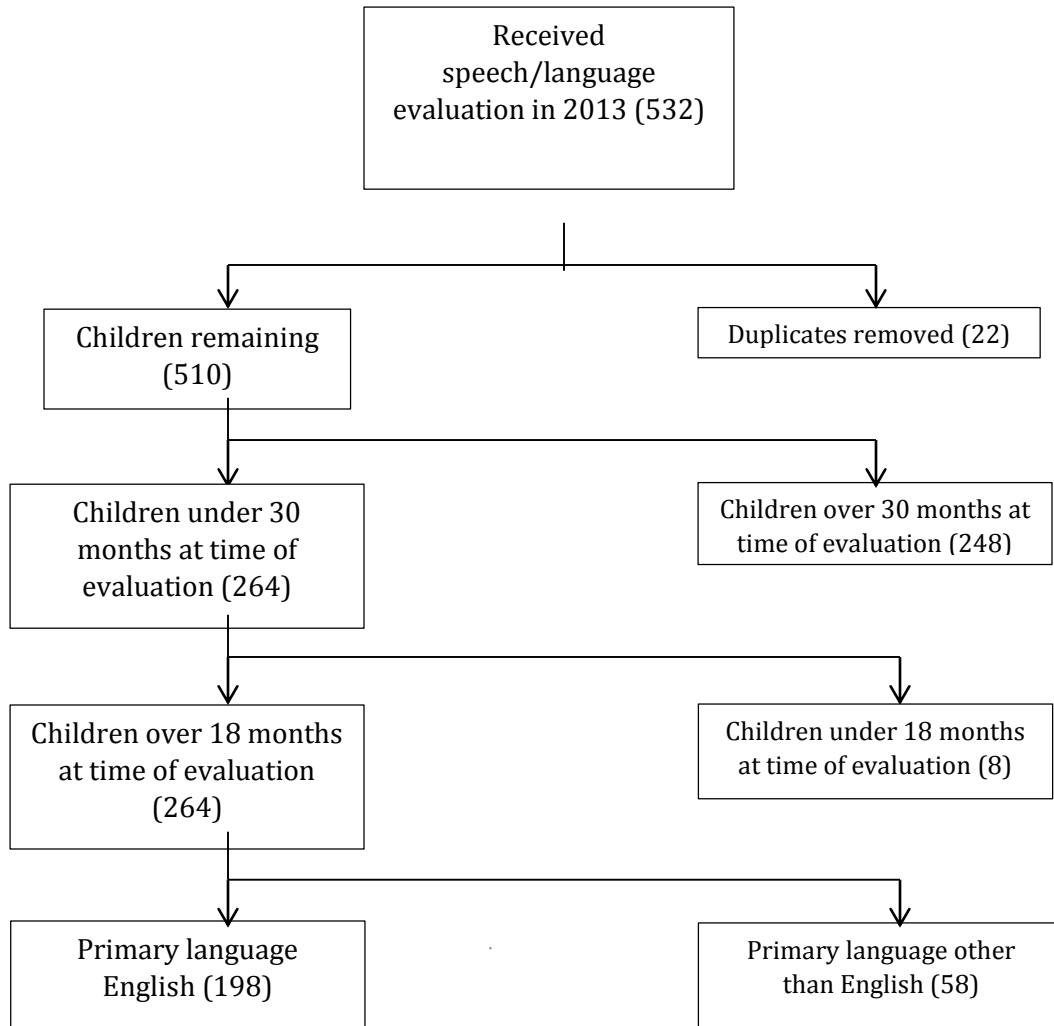


Figure 4. Hazard function for “Goals Met” group. Proportion of children who were discharged is on the Y axis by days of discharge (X axis).

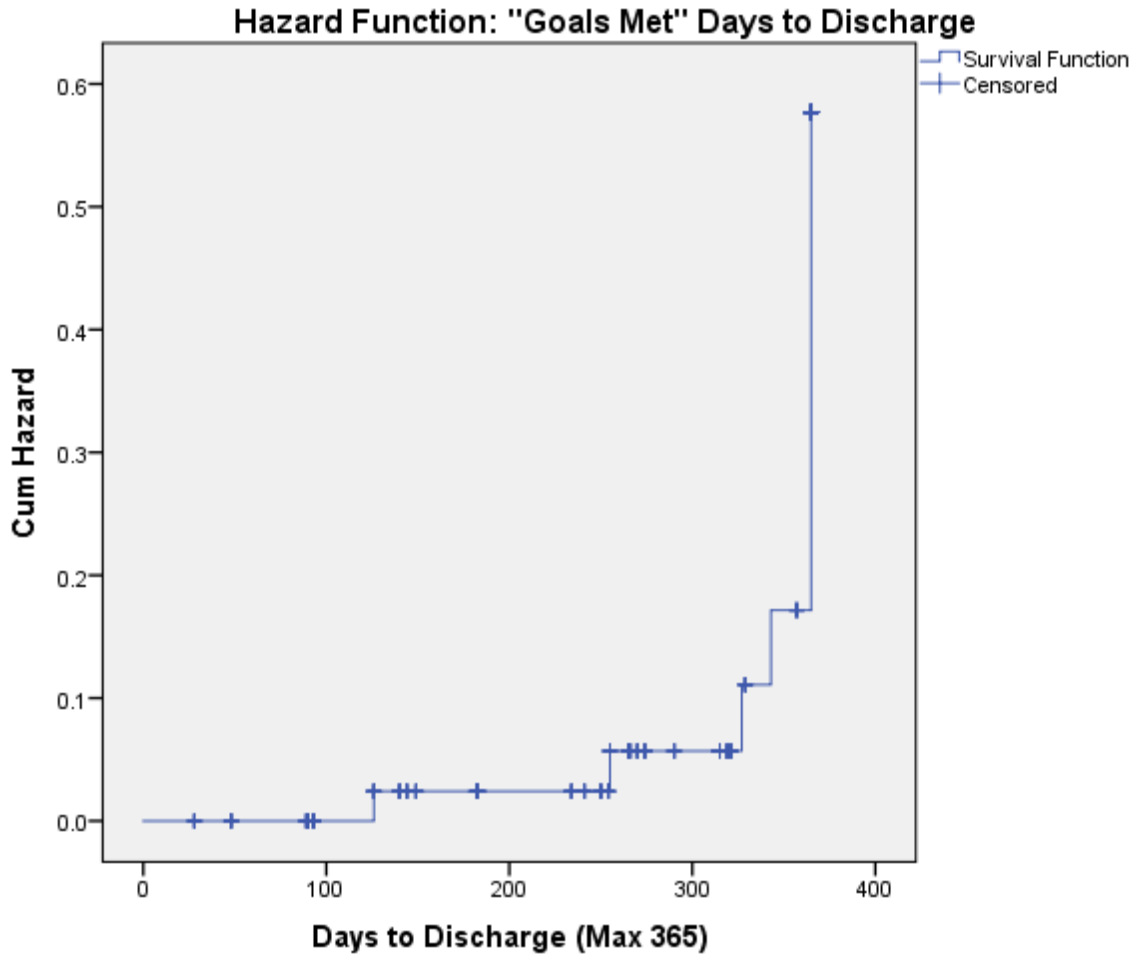


Figure 5. Hazard function for “Non-Attender” group. Proportion of children who were dismissed is on the Y axis by days of discharge (X axis).

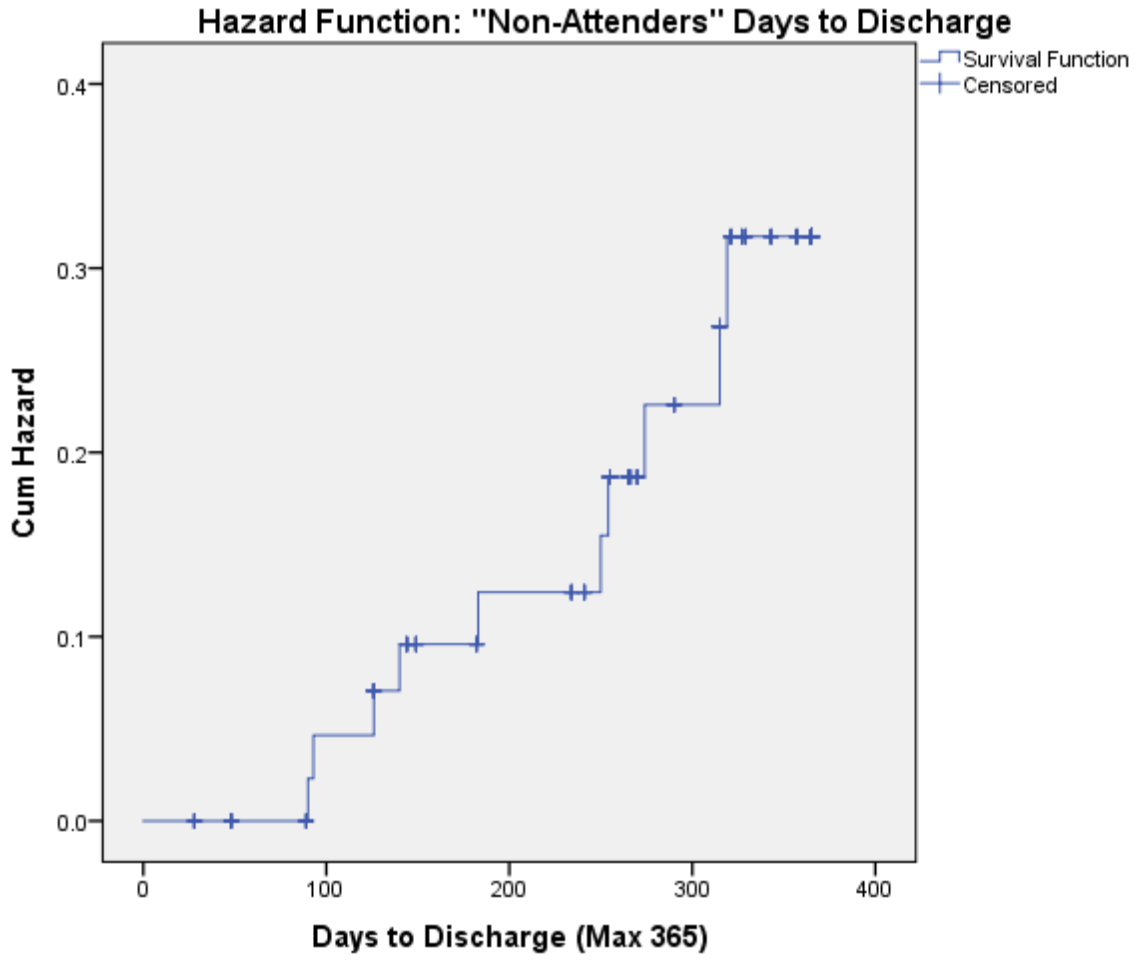


Figure 6. Hazard function for “Discharge to Schools” group. Proportion of children who were discharged is on the Y axis by days of discharge (X axis).

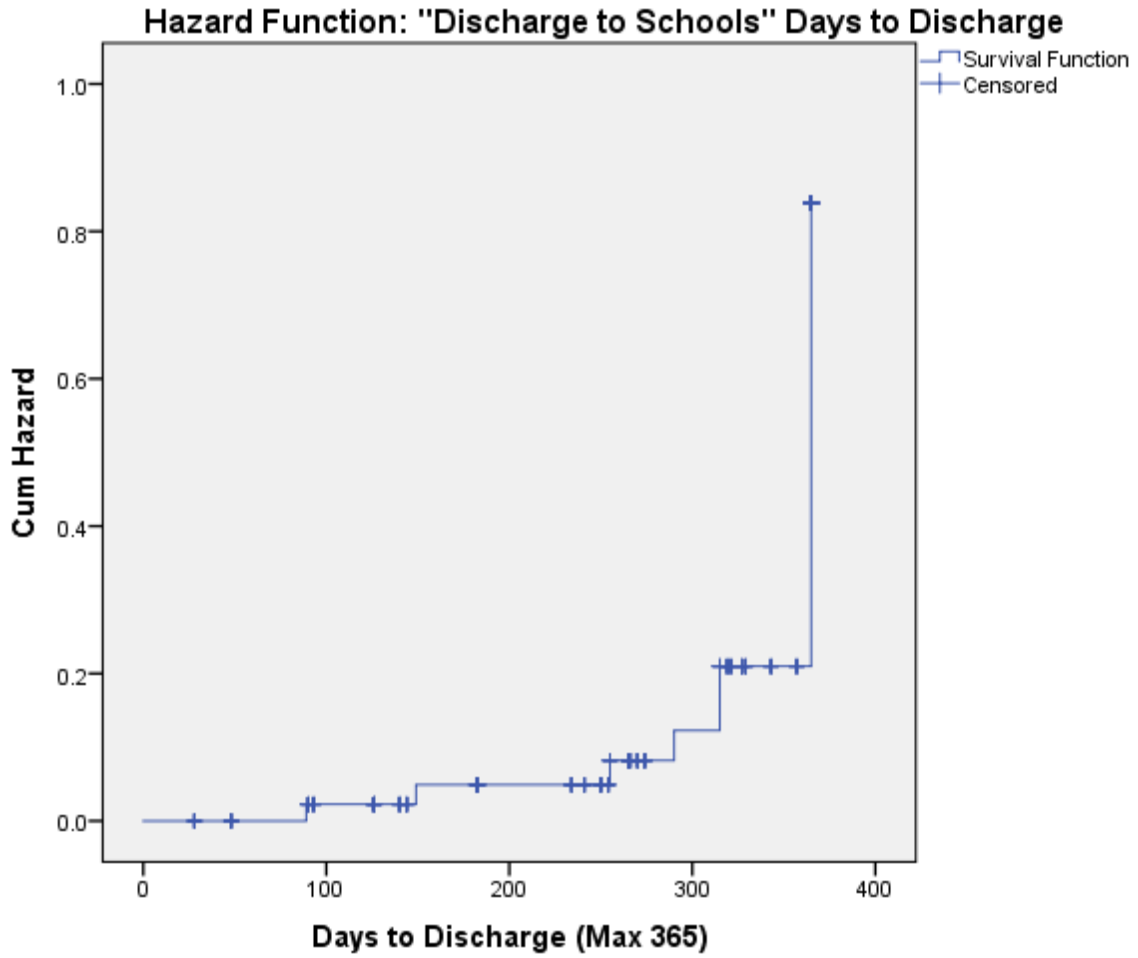
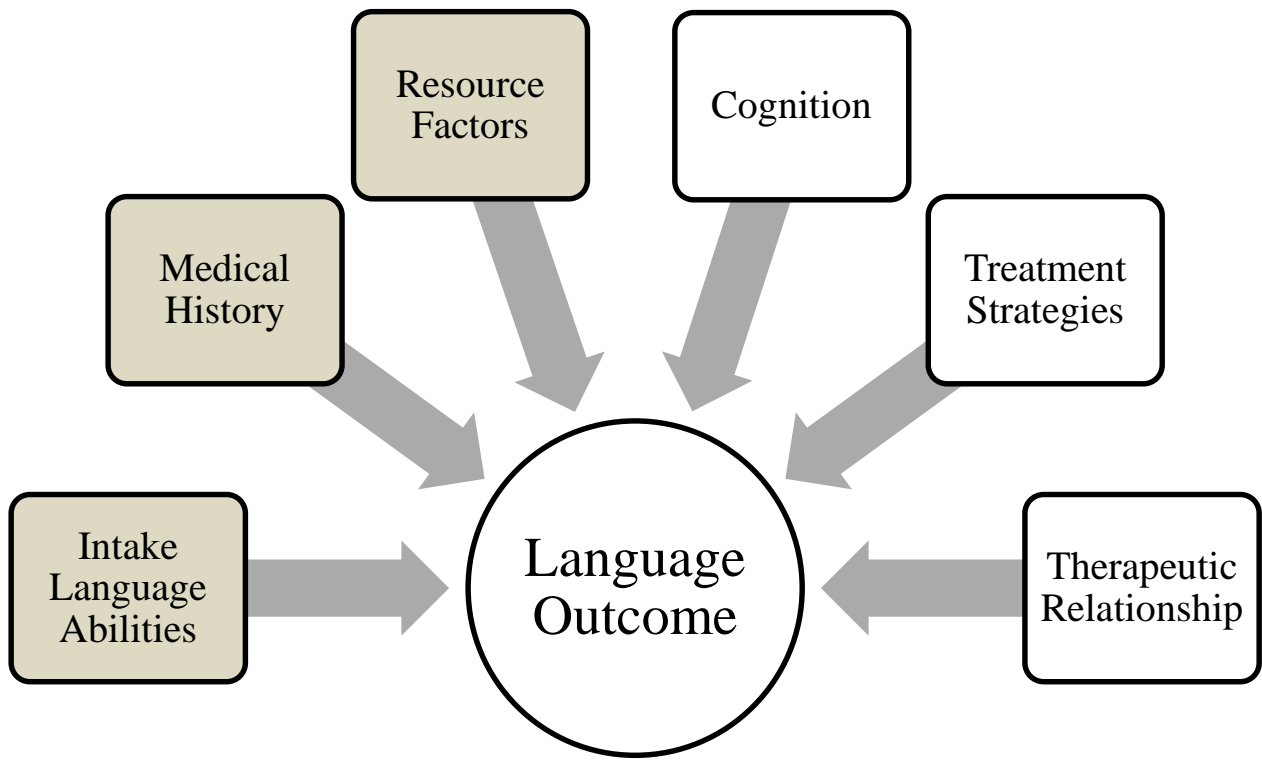


Figure 7. Prospective data analysis theoretical model. Highlighted input variables could be measured using retrospective data analysis; future work could include (some of) the additional variables.



APPENDICES

Appendix A: Non-English Speaking Children

Demographics. Children from non-English speaking homes account for 24% of the children evaluated at VBWC within the age range specified (n=58). All children were seen with the use of a trained medical interpreter or with video interpretation. Of these children, 58% came from Spanish-speaking homes. An additional 22% came from Arabic-speaking homes. Other languages included: American Sign Language, Bangla, Burmese, Chu, Gujarati, Hindi, Karen, Kurdish, Nepali, Persian, Russian, Uzbeki, and Vietnamese.

Referral to Treatment. Children from non-English speaking homes appear less likely to be referred directly to treatment and more likely to be referred for follow-up. Only two children (1%) in the English speaking sample were referred for follow-up evaluations. Of these two children, one did not return to Vanderbilt. The second was reassessed 8 months later and began treatment at that time. In contrast, 9 children (16%) of children who spoke languages other than English were referred for follow up (mean = 5 months, range = 3-9 months). In addition, 20 children were referred to daycare or preschool programs (34%), a recommendation that was made to only four (2%) of English-speaking children.

One possible reason for the recommendation of follow-up is that clinicians feel uncomfortable assessing through an interpreter. It is the interpreter who established rapport with the child and with the parents; the clinician may not be able to identify poorly-articulated words or phrases as they might in English; the clinician may not feel that they have gotten a full picture of the child's knowledge both in their native language and in English. (Current standards are to assess word knowledge and phrase structure in both language and use the sum, i.e., Bedore and Pena, 2008).

Standard Scores. The alternate explanation is that clinicians rely heavily on standard scores, particularly when they are not proficient in the language of testing. Standard scores were available for only 16 children (28%), all of whom received the Spanish Edition of the PLS-5 administered by a trained medical interpreter and supervised by a licensed speech pathologist. The mean receptive score on the PLS-5 was 80.94 (15.27) and the mean expressive score was 84.67 (9.60). There are not significant differences in scores between children who received the PLS-5 in English compared to children who received the PLS-5 in Spanish. However, note the discrepancy between scores on the PLS and scores on parent-reported measures in Table 11. If the PLS-5 lacks sensitivity, Spanish-speaking children are being under-identified for treatment in their native language. This is particularly important if English-speaking SLPs are relying on standard scores to determine treatment recommendations when they are not proficient in the language of the child's home.

Treatment. Treatment was recommended for 32 children (55%). 9 children (15%) enrolled in treatment, and only 5 children attended at least five sessions (8%). These numbers are significantly below the number of English speaking children who were recommended treatment (78%), enrolled in treatment (43%), and attended at least five sessions (34%). Because only five children attended treatment at VBWC, we cannot look at outcomes compared to English-speaking children. In addition, there are outcome scores for only two children (both Spanish-speaking). Of the five children, three remained in treatment after one year (all three continued to be enrolled on treatment for a full second year), and two children were discharged due to parent request. (The two children who were discharged due to parent request received services with different clinicians).

Appendix B: ICD-9 Codes Found in the Medical Charts

299 Autistic Disorder

312.9 Unspecified Disturbance of Conduct

*315.31 Expressive Language Disorder

*315.32 Mixed Expressive Receptive Language Disorder

*315.34 Speech and Language Developmental Delay Due to Hearing Loss

*315.39 Other Developmental Disorders of Speech and Language

315.5 Mixed Developmental Disorder

315.8 Other Specific Delays in Development

315.9 Unspecified Delay in Development

343.9 Cerebral Palsy

345 Epilepsy and Recurrent Seizures

389.0 Conductive Hearing Loss

389.1 Sensorineural Hearing Loss

389.12 Neural Hearing Loss

389.9 Unspecified Hearing Loss

426.7 (Heart Condition)

518.81 Acute Respiratory Failure

519 Tracheostomy

742.1 Microcephalus

742.2 Congenital Reduction Deformations of Brain

742.3 Congenital Hydrocephalus

749 Cleft Palate

749.2 Cleft Palate with Cleft Lip

756.0 Anomalies of Skull and Facial Bones

756.17 Spina Bifida

758.0 Down's Syndrome

758.32 Velo-Cardio-Facial Syndrome

758.39 Other Autosomal Deletions

758.6 Turner Syndrome

758.9 Unspecified Chromosomal Abnormality

759.83 Fragile X Syndrome

765.09 Extreme Immaturity

765.20 Unspecified Weeks of Gestation

770.89 Other Respiratory Problems After Birth

780.32 Complex Febrile Convulsions

780.39 Other Convulsions

783.42 Delayed Milestones

*784.51 Dysarthria and Anarthria

*Determined during speech/language evaluation