Image-Guided Cochlear Implant Programming: Pediatric Speech, Language, and Literacy

René Gifford, PhD
Stephen Camarata, PhD
Vanderbilt University Medical Center

STATISTICAL ANALYSIS PLAN AND POWER ANALYSES

Data Management. Dr. Dietrich (Co-I and biostatistician) will provide support for statistical analyses. All data will be stored in REDCap offering a secure, web-based application.

Scientific Rigor and Reproducibility. All data analyses and data sharing will adhere to the NIH's commitment to rigorous and transparent research. This will be accomplished through the analytic approach described here, which replicates our previous analytical approaches used for studies of adult IGCIP (24, 72) and our preliminary study of pediatric IGCIP (4). To achieve transparency, details will be reported that allow other research teams to reproduce the results. Furthermore, raw data will be presented in tables and appendices of our publications and will be made available upon request (within the scope and limits of IRB approved data sharing).

Statistical Analysis. *Overall strategy.* Statistical software (SPSS, STATA, R) will be used for the quantitative summarization of data and to test study hypotheses. The reliability of each of the scores from the standardized measures will be assessed and evaluated using Cronbach's alpha statistics. All analyses will be done using *intent-to-treat* principles. Statistical significance tests will maintain Type I error rates of no more than 0.05. Descriptive statistics will summarize and inspect the distributions of study measures for choosing the appropriate modeling procedure for testing hypotheses. A summary of aims, hypotheses, and associated statistical models are shown below in **Table 4**.

Missing data. Randomly missing responses to items within assessment tools will be handled via protocols specified by the instrument developers. When there is no protocol, if the participant has completed 75% or more of the items on a particular instrument, the mean score for that instrument will be calculated using available item responses and used in subsequent analyses. In-depth investigations of patterns of missing data will be undertaken to assess if data are missing due to random influences or if there are certain study conditions (e.g. waitlist control) or participant characteristics (e.g., age, hearing function) that are more or less likely to be associated with certain patterns of missing data (i.e. lost to follow-up). We expect that most assessments will not be missing at randomization, thus imputation would not be appropriate.

TABLE 4

HYPOTHESES	MEASURES	STATISTICAL MODELS
<u>AIM 1</u> • significant positive short-term gain in auditory function for children receiving IGCIP <u>AIM 2</u> • differential growth in auditory function will predict growth in PA, which will predict mediated growth for reading measures <u>AIM 3:</u> • significant positive growth in speech & language; this growth will be predicted by the relative improvement in auditory function from the IGCIP	<u>AIM 1</u> • Spectral, temporal, & spectro-temporal resolution • Speech recognition • Subjective questionnaires (auditory & quality of life) <u>AIM 2</u> • PA (Tests & Tasks) • Reading outcomes (Tests) • Control for working memory (Tests & Tasks) <u>AIM 3</u> • Language (Tests of expressive, receptive, and narrative) • Speech production (Tests and Acoustic Analysis)	<u>AIMS 1-3</u> • **Descriptive statistics** of all measures • **Bootstrapped 95% confidence intervals for all effects** • **Statistical Control of Potential Confounds:** Covary baseline levels of nonverbal cognition, working memory, and speech recognition • **Mixed-effects modeling:** Test the differential effect of IGCIP on the trajectories of change in auditory function. 30/group. Minimum detectable effect size = 0.67 (SDs at endpoint); traditional Cohen effect sizes d = 0.2/0.5/0.8 ~ small/medium/large. • **Cross-lagged panel and path analysis:** Test the mediation effect of key factors on reading outcomes (e.g., phonological awareness on the relationship between auditory function/speech recognition on reading ability and speech production). Minimum detectable path coefficient 0.35 (~12% shared variance) • **z-test of independent path coefficients.** Test for differences in the size of the path coefficients between the two study groups.

Aim 1 and Aim 2: Analysis & hypotheses testing: The outcome variables are auditory function, speech recognition, PA, working memory, and reading gains over various time points (**Approach, Table 3**). Descriptive and graphical summaries of trajectories by study group will be conducted initially for detection of outliers and to provide insight into patterns of change. Key statistical tests will involve study group (IGCIP vs. waitlist control) comparisons of the mean slopes resulting from differences in baseline and post-intervention assessments. Tests will be conducted using general linear mixed or multilevel analysis. While randomization ensures equal opportunity for study conditions, it does not ensure equivalence of baseline values. If it is found that group baseline values differ, baseline scores will be included as a covariate in the analysis as will potential confounds such as baseline intellectual level and working memory ability. Within this general multilevel statistical approach, hypothesized differences will be tested by assessing the statistical significance of the main and interaction effects of study group on time-related contrast in baseline and study assessment points in the outcome variable scores. In other words, we expect that the slope of outcome measure scores in the waitlist control group will be nearly '0' while those of immediate IGCIP group will demonstrate a statistically significant positive slope. In addition to statistical significance testing, bootstrapping methods will be used to generate 95% confidence intervals for all sample descriptive (e.g., group means at each time of assessment) and effect estimates (e.g., eta-squared for group effect on linear slope of the outcome scores). Because we expect there to be correlations amongst the multiple outcome measures, a multivariate approach will provide more unified (systemic) statistical test of the intervention effects.

Aim 3: Analysis and hypothesis testing: The focus of this aim is to explore the complex relationships among changes in the various measures of hearing, speech, and language. As an example of this approach, **Figure 7** displays an example cross-lagged panel analysis which illustrates the structure for statistical analysis of this aim. Via comparisons of the strength of the relationships between the changes in one domain at time 1 with the changes in another domain at time 2, etc., this type of analysis will maximize the information gained from the longitudinal assessment of the multiple domains and inform additional causal hypotheses for subsequent research. Bootstrapped 95% confidence intervals will be generated for each of the path coefficients. For all statistical analyses, we will allow for covariates associated with the child and family including chronological age at assessment, age at CI, age at identification, nonverbal cognition, working memory, gender, and socioeconomic status (10, 15, 119, 120).

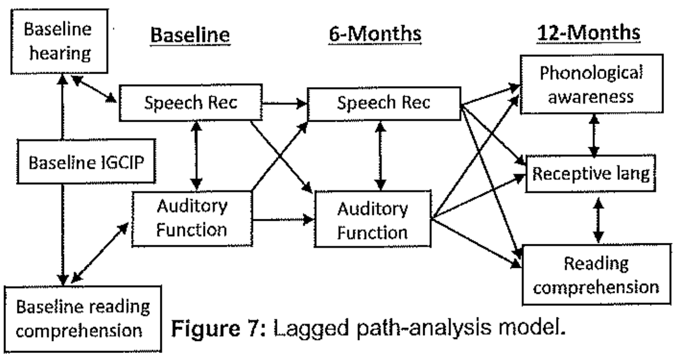

Figure 7: Lagged path-analysis model.

Sample size and power. Sample size estimates are based on the desire to detect clinically meaningful effects of the intervention using information from our preliminary studies while maintaining study feasibility. An analysis sample of 30 participants per study group will provide 80% statistical power (two-sided α=0.05) for the detection of an intervention effect on the trajectories of the hearing, speech, language, PA and reading as small as 0.32 (Cohen's *d* equivalent=0.67, adjusted for baseline with $\eta^2 \geq 0.2$) and 0.35 (Cohen's *d* equivalent =0.74, unadjusted). Differences of this magnitude are considered to be clinically meaningful. Furthermore, the statistical power estimates are conservative due to the proposed used of mixed-effects analyses approaches that will enable the increased power of treating the repeated assessments as independent values yet appropriately adjusting the standard errors for the correlations among those repeated assessments. The proposed final sample of 60 will enable detection of a path correlation as small as 0.35 (80% power, 2-tailed α=0.05). Correlational values of that magnitude or larger were observed in our preliminary work. Detectable differences between the strength of two path coefficients will be 0.4-0.5 (80% power, 2-tailed α=0.05) depending on the value of the coefficients and the size each correlation has with other values. The focus of the cross-lagged panel analysis will be on generating effect sizes deepening our understanding of the mechanisms underlying effects of change in hearing on higher-level PA/speech/language downstream. Accounting for 20% attrition, we will enroll 72 patients to achieve a 60-subject sample.