PHARMACEUTICAL INTERVENTIONS FOR HEARING LOSS (PIHL)

Reference the following open access material as:

Slade, Martin. Pharmaceutical Interventions for Hearing Loss. (Oct 2014). *Statistical Considerations* [Guidelines]. Available from http://hearing.health.mil/EducationAdvocacy/Newsletters.aspx

Statistical Considerations

Martin D. Slade, MPH

There are several statistical issues to consider when planning a study to measure hearing loss. These range from the learning curve associated with audiometric testing to the fact that multiple measurements of hearing level are recorded on a subject each time that they undergo a hearing evaluation. The following sections describe statistical issues that need to be considered, often recommending methodologies to utilize.

Audiometric Learning Curve
Literature suggests that individuals new to audiometric testing experience a learning curve during which time their measured hearing improves. This usually occurs over the first 4 tests (Royster & Royster, 1986). Due to this effect, an individual's first 4 audiometric tests should be discounted when calculating changes in threshold hearing levels. One potential solution to this issue is to have individual's that have had less than 4 audiometric tests to undergo several rounds of audiometric testing prior to recording a baseline threshold level.

Definition of Hearing Loss / Multiple Comparisons
Perhaps the greatest issue with regard to studies involving hearing loss as an outcome is the fact that threshold hearing levels are measured at many different frequencies. For audiometric testing, these typically include 0.5, 1, 2, 3, 4, 6, and 8k Hz. Therefore, seven separate hearing threshold levels are recorded. Analyses conducted to determine if there is an effect of an intervention at any of these frequencies result in seven statistical tests being performed. Thus, to be able to show an effect with 95% confidence (α=0.05), utilizing the Bonferroni correction for multiple comparisons, the observed probability (p-value) would need to be below 0.05/7 = 0.0071. Given that there were

seven frequencies included in this analysis meant that we were taking an average of both ears. If, instead, we were looking at right and left ears separately, then there would be 14 different statistical tests performed and the required p-value (to demonstrate an effect with 95% confidence) would need to be below 0.05/14 = 0.0036. Any additional definitions of hearing loss such as averages of certain frequencies would only add to the number of statistical tests performed thereby making the required p-value even harder to demonstrate.

There are two different approaches to reducing the multiple comparisons problem. The first of these is to define one specific definition of hearing loss to be used as the primary objective. Thus, a definition such as the average of frequencies 2, 3, and 4k Hz across both ears would eliminate the entire multiple comparison issue. Other definitions of hearing loss can surely be evaluated, but they should be explicitly denoted as secondary analyses. These should be considered as exploratory analyses and would, therefore, be reported without correcting for multiple comparisons. Note, however, that any significant predictors will need to be replicated in a follow-up trial whereby this definition of hearing loss becomes the primary outcome.

A second approach to reducing the multiple comparisons problem would be to substantially increase the sample size such that the collected data could be randomly split. The first group would be used in an exploratory analysis whereby various definitions of hearing loss could be evaluated. Once that is completed and the desired definition of hearing loss determined, the second group would be used to confirm the results. This second analysis would also be the one that is reported (Muller, Otto & Benignus, 1983).

Some researchers have suggested that a repeated measures ANOVA be utilized as a way to remove the problem of multiple comparisons. Unfortunately, without interactions, the model assumes that the intervention affects not only both ears, but also all of the frequencies, in a like manner. If, however, only certain frequencies are actually affected, this method will lose substantial power. This limitation can be overcome by including interactions, but this will substantially decrease the power of the study.

Along with continuous definitions of hearing loss, binary definitions of hearing shifts should also be considered. In choosing this definition the amount of hearing loss expected to occur during the trial needs to be taken into account such that there are enough expected cases (shifts). This should be determined from power calculations to determine sample size requirements.

Independence/Dependence of Left and Right Ear Test Results

There has been previous discussion as to whether an analysis that utilizes a subject's right ear results separately from their left ear results should be considered independent. In other words, should they count as two independent tests? The simple answer is that these results are NOT independent. They are both occurring within a single individual, so if there are genetic effects that related to hearing loss, then that person's genetics are affecting both these measurements. The modeling of this type of data should incorporate a random effect of individual to control for the covariance of the measurements.

References

Royster, J. D., & Royster, L. H. (1986). Using audiometric data base analysis. *Journal of Occupational and Environmental Medicine*, *28*(10), 1055-1068.

Muller, K. E., Otto, D. A., & Benignus, V. A. (1983). Design and analysis issues and strategies in psychophysiological research. *Psychophysiology*, *20*(2), 212-218.

http://hearing.health.mil/EducationAdvocacy/Newsletters.aspx

The Pharmaceutical Interventions for Hearing Loss (PIHL) Newsletter is published by the Department of Defense Hearing Center of Excellence. This Newsletter may include information that was obtained from publicly available sources. The views expressed represent the personal views of the authors and do not imply Department of the Air Force or DOD endorsement of factual accuracy or opinion. The information is presented for information purposes only. While this information has been gathered from reliable sources, its currency and completeness cannot be guaranteed. Your comments are wecome:

HCE Email: sharon.bryant.4.ctr@us.af.mil, Phone: 210-292-4100. COL Mark Packer, Director. Tanisha Hammill, Editor, Email: tanisha.hammill.ctr@us.af.mil, Phone: 210-292-5641