PHARMACEUTICAL INTERVENTIONS FOR HEARING LOSS (PIHL)

Supra-threshold Testing Using Speech-in-Noise and Auditory Evoked Potentials

Colleen G Le Prell PhD, Douglas S Brungart PhD

In this white paper, we discuss the use of supra-threshold tests, including supra-threshold auditory brainstem response (ABR) and speech-in-noise testing, for early detection of noise-induced hearing loss (NIHL) and for assessing the effectiveness of interventions designed to prevent NIHL. In humans, the accepted clinical standard for detecting NIHL is the behavioral audiogram, which is based on the absolute detection threshold of narrow-band noises or tones (for detailed review, see Le Prell & Lobarinas, 2015). There is little question that the audiogram is a reliable and sensitive measure of NIHL in human listeners. However, recent results from noise exposed animals suggest that NIHL can cause substantial neurodegeneration in the peripheral auditory system without degrading the pure tone audiometric threshold, as measured by the threshold ABR (Le Prell & Bao, 2012; Kujawa &Liberman, 2009). This suggests that clinical measures of auditory performance that are conducted with stimuli presented above the detection threshold, such as speech-in-noise testing or the supra-threshold ABR, may be more sensitive than the behavioral audiogram in detecting early stage NIHL in listeners with audiometric thresholds within normal limits.   Also, from a more practical standpoint, it is often difficult to find locations with low enough ambient noise levels for the collection of valid behavioral audiometric threshold data in the austere working environments where occupational noise exposures are most likely to occur. This can cause logistical difficulties in the design of studies aimed at the detection of small changes in the hearing abilities of noise exposed listeners. These environmental concerns can provide additional motivation to use supra-threshold speech-in-noise testing or supra-threshold

ABR responses as a supplement to the behavioral audiogram in the measurement of small changes in hearing loss in noise-exposed listeners.

Speech-Based Functional Tests

Individuals with high-frequency hearing loss are impaired in terms of their ability to understand speech in noise because some parts of the speech signal are inaudible due to their elevated thresholds (Quist-Hanssen et al., 1979; Badri et al., 2011). However, most hearing scientists believe that, in addition to this *audibility* component, that there is also a *distortion* component of hearing loss that causes hearing impaired listeners to perform poorly on tests of speech-in-noise perception even when all components of the speech signal are theoretically audible (Plomp, 1986). Consequently, the American Academy of Otolaryngology (AAO) recently recommended word recognition test scores be collected in *all* clinical trials that assess auditory function (Gurgel et al., 2012). Although the absolute nature of this guidance has been questioned by the American Academy of Audiology, which believes that speech-in-noise testing, specifically, is not optimal for *all* clinical trial populations (Carlson, 2013), there are many reasons to believe that tests using speech- in-noise backgrounds may play an important role in providing a "stress test" for auditory function (Wilson, 2011) that may be more sensitive to small changes in hearing ability than a behavioral audiogram. Moreover, interest and enthusiasm for speech-in-noise tasks as a proxy for auditory function should be high given suggestions that loss of neural connections from inner hair cells, resulting in decreased ABR amplitude *in the absence of overt threshold shift,* may underlie speech-in-noise discrimination deficits (Kujawa &Liberman, 2009; Lin et al., 2011; Makary et al., 2011).

The impact that interfering noise has on speech perception has long been a subject of interest to hearing scientists, and over the past 70 years a substantial number of speech-in-noise tests have emerged in the literature. Most of these tests were initially designed for use in hearing research (e.g., Carhart & Tillman, 1970), although many of them have also been adapted for use in the clinic (Carhart, 1951, for review, see Wilson, 2011). Early example of tests that were initially designed primarily for research have included the Speech Perception in Noise (SPIN) Test (Kalikow et al., 1977), the Revised Spin Test (R-SPIN) (Bilger et al., 1984), the Connected Speech Test (CST), (Cox et al., 1987; Cox et al., 1988), the Speech Intelligibility Rating (SIR) test (McDaniel & Cox, 1992; Beck & Speaks, 1993), and the Revised Speech Intelligibility Rating (RSIR) test (Speaks et al., 1994).

In recent years, a few speech-in-noise tests have begun to be used more widely in the clinic. Two examples are the Hearing in Noise Test (HINT) as modified for American English (Nilsson et al., 1994) and the Speech Recognition in Noise Test (SPRINT) (AR 40-501; Standards of Medical Fitness, http://www.apd.army.mil/pdffiles/r40_501.pdf), both of which are used primarily for evaluating the fitness for duty of hearing impaired individuals in high-risk occupations such as police and fire departments and in the military. Three other widely-used speech-in-noise tests include the Quick Speech-In-Noise (QuickSIN) test (Killion et al., 2004), the Words-in-Noise (WIN) test (for review, see

Wilson, 2011), and the Oldenburg Matrix Sentence Tests (OLSA) (Wagener, Brand, & Kollmeier, 1999), which is used primarily in Europe. These tests are used to assess general speech-in-noise performance and, in some cases, to assess hearing aid or cochlear implant candidacy. It is difficult to provide good (evidence-based) guidance on the selection of speech-in-noise tasks for clinical trials, as there have not been many studies that provide empirical data directly comparing performance of participants across speech-in-noise tests, although there are some notable examples of studies that do offer back-to-back comparisons, such as Wilson et al. (2007) and Grant and Walden (2013). One finding of these back-to-back comparisons is that the WIN tends to be one of the more challenging tests for young listeners with normal hearing, which may be helpful in ensuring that no floor effects occur in studies designed to detect small changes in hearing in this population.

In general, the correlation between speech-in-noise scores and the behavioral audiogram is not very high, suggesting that factors other than pure audibility have a strong influence on speech perception performance in hearing impaired listeners (Smorenberg, 1992). However, recent research has shown certain speech-in-noise tests, and in particular those that adaptively determine the speech reception threshold of three-digit sequences presented in noise, can be reasonably effective in identifying individuals with abnormal hearing (i.e., Pure Tone Averages at 0.5, 1.0, and 2.0 kHz >20 dB), even when administered over a relatively crude telephone line (Watson, Kidd, Miller, Smits, & Humes, 2012).

In terms of their utility for assessing small changes in hearing in a clinical evaluation of oto-protectants, there are potential advantages to the use of speech-in-noise tests as a supplement to the pure tone audiogram. Speech-in-noise stimuli can and should be presented well above threshold levels, meaning that they could potentially be conducted under noise-attenuating headphones in office or clinic spaces with background noise levels that would be too high to obtain a behavioral audiogram. They also provide an objective measure of performance that allows listeners to be provided with correct answer feedback and may be more stable over time than the subjective judgment of the presence or absence of a low-level tone. Finally, they provide a stimulus with sufficient ecological validity to independently sustain a claim of efficacy in a clinical trial focused on the elimination of NIHL: few would question the utility of an intervention that preserved speech-in-noise performance even if it had no impact on the pure tone audiogram. However, it must also be recognized that studies focused on the effectiveness of NIHL interventions are implicitly based on sequential changes in metrics that occur within a single subject pre- and post-noise exposure, and that there is very little data currently available to suggest how effective speech-in-noise tests might be when used in this fashion.

One fundamental issue that must be addressed is the possibility of learning effects across repeated administrations of the same speech-in-noise test. Most tests with open-set speech materials have at least a few different lists of words that are purported to be "equivalently difficult", but when using one of these tests the experimenter always has

to make a trade-off between the number of times the test is administered and the number of lists to use per test (to improve reliability). Closed-set tests, such as the OLSA or the Triple-Digit test may be more suitable for studies evaluating longitudinal changes in hearing loss.

Also, although it is well known that speech-in-noise performance degrades with increasing hearing loss, the temporal relationship between speech-in-noise performance and hearing loss is not known. It may be the case that small changes in the speech-in-noise score will be seen before changes in the behavioral audiogram are detectable, or it may be that changes in the audiogram will always occur before a change in speech in noise performance.   Thus, at this point, it must be admitted that there is some risk that speech-in-noise testing will not be as sensitive as the audiogram for detecting changes in hearing sensitivity. Nevertheless, the inclusion of speech-in-noise testing where possible is recommended, as the potential benefits of finding small changes in hearing that are not detectable on the behavioral audiogram are likely to outweigh the relatively modest costs of including a short speech-in-noise test battery as part of the pre- and post- noise exposure evaluation.

Auditory Brainstem Response (ABR)
ABR supra-threshold input-output functions are considered to be a "gold standard" metric for assessing lasting effects of noise on the neural population in animal testing. Interest in ABR tests in humans has been increasing given reported loss of neural connections from inner hair cells, with corresponding decrease in Wave I ABR amplitude, after noise exposures that produce robust threshold shift (i.e., approximately 40-dB temporary threshold shift, or "TTS" (see "Temporary and Permanent Noise-Induced Threshold Shifts" guidance document for further definition), measured 24-hours post-noise) in rodents (Kujawa & Liberman, 2009; Lin et al., 2011; Wang & Ren, 2012). Limited physiological data from humans have not shown corresponding deficits however. There were no deficits in ABR amplitude in either Veterans with known noise exposure (Konrad-Martin et al., 2012) or professional pop/rock musicians (Samelli et al., 2012) compared to control subjects. Additional controlled studies are needed. If human ABR amplitude varies with noise exposure history, in the absence of threshold deficits, we may ultimately gain insight into the "critical boundary" at which noise becomes hazardous to synapses in human ears, a boundary that is not currently known for either humans or laboratory animals (for discussion of critical boundary, see Le Prell et al., 2012; Spankovich et al., 2014). One problem in human testing of the ABR is that the absolute level of the ABR can vary according to the placement and sensitivity of the electrodes used to make the measurement. If an attempt to measure Wave I ABR recordings is attempted as part of a clinical trial, every effort must be made to increase the sensitivity of the measure. Possible strategies include using electrodes that utilize the ear canal as a recording site to improve sensitivity (Gaddam & Ferraro, 2008), and normalizing the electrode sensitivity by using the of Wave I/ Wave V amplitude ratio rather than an absolute measure of Wave I output (Musiek et al., 1984). Amplitude of Wave I may not be the only evoked potential of interest. Recently, Ruggles et al. (2011;

2012) identified variability in the auditory brainstem frequency-following response in normal-hearing young adults as a key factor for difficulty discriminating sounds in noise. Other groups are independently arriving at similar conclusions; Hopkins and Moore (2011) also reported reduced sensitivity to temporal fine structure cues may underlie speech perception difficulties.

Challenges to the field
Oto-protection studies have focused on the ability of agents to reduce or prevent threshold shifts. However, an emerging body of evidence suggests that changes in supra-threshold hearing may also be important. Provocative data from animal studies suggest that robust TTS can be accompanied by damage to auditory nerve fiber synapses. Deficits on speech in noise tasks are common, but there is little agreement which tests to use to document deficits. Noise-induced ABR amplitude changes are a "hot-topic" in animal models, but such changes have not been established in humans. This is clearly relevant to long-term trajectory of noise-induced changes in hearing. Slow-to-develop loss of the spiral ganglion fibers and increased changes in hearing associated with aging were observed over the course of the mouse life span after robust TTS (Kujawa & Liberman, 2006). Those data contradict a long-standing assumption that, as long as thresholds recover, the ear has completely recovered. These reports have raised important questions, such as whether a more moderate TTS could also induce long-term deficits through similar neural-based mechanisms. The "critical boundary" at which long-term synaptic deficits are first induced is not known in rodents or humans (for discussion, see Le Prell et al., 2012; Spankovich et al., 2014), and is subject to some debate. Some evidence suggests the phenomena is likely in humans (Plack et al., 2014). In contrast, the Institute of Medicine (2005) report on Noise and Military Service specifically concluded, "The committee's understanding of the mechanisms and processes involved in the recovery from noise exposure suggests, however, that a prolonged delay in the onset of noise-induced hearing loss is unlikely."

Given all the evidence suggesting that some functional changes in hearing can occur prior to the point where a significant threshold shift is seen in the audiogram, we believe that strong arguments can be made for the inclusion of some supra-threshold measure of performance as part of any large-scale study designed to detect small changes in hearing in noise-exposed populations, including those focused on evaluating the efficacy of interventions for noise-induced hearing loss. Unfortunately, because few if any studies to date have compared the sensitivity of different speech-in-noise tests to longitudinal, within-subject changes in hearing ability, it is difficult to define a single best practice for conducting such tests. If speech in noise testing is confined to just a small number of measurement sessions per subject, any of a number of clinical tests could be used for the pre- and post- testing, including the widely-available WIN and QuickSIN. However, if a larger number of periodic test sessions are required, it may be desirable to switch to a closed set test such as the OLSA or the Triple Digit test, which can be scored automatically and can be administered as many times as necessary without significant learning effects. If time permits and it is possible to include a speech-in-noise test more

sensitive than the speech-in-noise tests typically included in a standard clinical battery, then it might be reasonable to include a more challenging speech perception condition where the target talker is presented over headphones at a simulated location in front of the listener and two flanking maskers speech maskers are located to the left and the right of the target talker. Examples of such tests include the modified version of the QuickSIN, where the target talker is time-compressed and presented in the presence of simulated room reverberation proposed by Brungart, Sheffield, and Kubli (2014); the modified version of the Coordinate Response Measure proposed by Gallun, Diedesch, and Campbell (2013); or the Listening in Spatialized Noise test (LISN-S), proposed by the National Acoustics Laboratory in Australia (Cameron, Glyde & Dillon, 2011). These tests have all been shown to be more sensitive than standard speech-in-noise tests to auditory deficits in listeners who complain about speech in noise problems but have relatively normal audiometric thresholds. Because there is so little information on longitudinal changes in speech-in-noise perception in noise exposed populations, we contend that the most important consideration right now is to make every effort to ensure that *some* speech-in-noise testing is included in the design of large-scale studies examining the effects of noise exposure on hearing ability. We believe the costs of adding an additional speech-in-noise measure would, in most cases, be relatively minor compared to the large-scale logistical costs of identifying and sequentially testing a large population of listeners pre- and post- noise exposure, and that the results of these studies will provide invaluable information that will help shape the design of future studies examining interventions for NIHL and answer basic questions about human hearing that have applications far beyond the results of any single study conducted in this area.

References

Badri R, Siegel JH, Wright BA (2011) Auditory filter shapes and high-frequency hearing in adults who have impaired speech in noise performance despite clinically normal audiogramsa. J Acoust Soc Am 129:852-863.

Beck WG, Speaks C (1993) Intelligibility of selected passages from the Speech Intelligibility Rating (SIR) test. J Speech Hear Res 36:1075-1082.

Bilger RC, Nuetzel JM, Rabinowitz WM, Rzeczkowski C (1984) Standardization of a test of speech perception in noise. J Speech Hear Res 27:32-48.

Brungart, D. S., Sheffield, B. M., & Kubli, L. R. (2014). Development of a test battery for evaluating speech perception in complex listening environments. The Journal of the Acoustical Society of America, 136(2), 777-790.

Cameron, S., Glyde, H., & Dillon, H. (2011). Listening in Spatialized Noise—Sentences Test (LiSN-S): Normative and Retest Reliability Data for Adolescents and Adults up to 60 Years of Age. Journal of the American Academy of Audiology, 22(10), 697-709.

Carhart R (1951) Basic principles of speech audiometry. Acta Otolaryngol 40:62-71.

Carhart R, Tillman TW (1970) Interaction of competing speech signals with hearing losses. Arch Otolaryngol 91:273-279.

Carlson DL (2013) American Academy of Audiology response to Gurgel et al. Otolaryngol Head Neck Surg 149:349-350.

Cox RM, Alexander GC, Gilmore C (1987) Development of the Connected Speech Test (CST). Ear Hear 8:119S-126S.

Cox RM, Alexander GC, Gilmore C, Pusakulich KM (1988) Use of the Connected Speech Test (CST) with hearing-impaired listeners. Ear Hear 9:198-207.

Gaddam A, Ferraro JA (2008) ABR recordings in newborns using an ear canal electrode. Int J Audiol 47:499-504.

Gallun, F. J., Diedesch, A. C., Kampel, S. D., & Jakien, K. M. (2013). Independent impacts of age and hearing loss on spatial release in a complex auditory environment. Frontiers in neuroscience, 7.

Grant KW, Walden TC (2013) Understanding excessive SNR loss in hearing-impaired listeners. J Am Acad Audiol 24:258-273; quiz 337-258.

Gurgel RK, Jackler RK, Dobie RA, Popelka GR (2012) A new standardized format for reporting hearing outcome in clinical trials. Otolaryngol Head Neck Surg 147:803-807.

Hopkins K, Moore BC (2011) The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. J Acoust Soc Am 130:334-349.

Institute of Medicine (2005) Noise and Military Service: Implications for Hearing Loss and Tinnitus. Washington D. C.: The National Academies Press.

Kalikow DN, Stevens KN, Elliott LL (1977) Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. J Acoust Soc Am 61:1337-1351.

Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S (2004) Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J Acoust Soc Am 116:2395-2405.

Konrad-Martin D, Dille MF, McMillan G, Griest S, McDermott D, Fausti SA, Austin DF (2012) Age-related changes in the auditory brainstem response. J Am Acad Audiol 23:18-35; quiz 74-15.

Kujawa SG, Liberman MC (2006) Acceleration of age-related hearing loss by early noise exposure: evidence of a misspent youth. J Neurosci 26:2115-2123.

Kujawa SG, Liberman MC (2009) Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. J Neurosci 29:14077-14085.

Le Prell CG, Bao J (2012) Prevention of noise-induced hearing loss: potential therapeutic agents. In: Noise-Induced Hearing Loss: Scientific Advances, Springer Handbook of Auditory Research(Le Prell, C. G. et al., eds), pp 285-338 New York: Springer Science+Business Media, LLC.

Le Prell CG, Dell S, Hensley BN, Hall JWI, Campbell KCM, Antonelli PA, Green GE, Miller JM, Guire K (2012) Digital music exposure reliably induces temporary threshold shift (TTS) in normal hearing human subjects. Ear Hear 33:e44-58.

Le Prell CG, Lobarinas E (2015) Strategies for assessing antioxidant efficacy in clinical trials. In: Oxidative Stress in Applied Basic Research and Clinical Practice: Free Radicals in ENT Pathology(Miller, J. M. et al., eds) New York: Springer.

Lin HW, Furman AC, Kujawa SG, Liberman MC (2011) Primary neural degeneration in the guinea pig cochlea after reversible noise-induced threshold shift. J Assoc Res Otolaryngol 12:605-616.

Makary CA, Shin J, Kujawa SG, Liberman MC, Merchant SN (2011) Age-related primary cochlear neuronal degeneration in human temporal bones. J Assoc Res Otolaryngol 12:711-717.

McDaniel DM, Cox RM (1992) Evaluation of the speech intelligibility rating (SIR) test for hearing aid comparisons. J Speech Hear Res 35:686-693.

Musiek, F. E., Kibbe, K., Rackliffe, L., & Weider, D. J. (1984). The auditory brain stem response IV amplitude ratio in normal, cochlear, and retrocochlear ears.*Ear and hearing*, 5(1), 52-55.

Nilsson M, Soli SD, Sullivan JA (1994) Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am 95:1085-1099.

Plack CJ, Barker D, Prendergast G (2014) Perceptual consequences of "hidden" hearing loss Trends in Hearing 18.

Plomp, R. (1986). A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. *Journal of Speech and Hearing Research, 29*(2), 146-154.

Quist-Hanssen S, Thorud E, Aasand G (1979) Noise-induced hearing loss and the comprehension of speech in noise. Acta oto-laryngologica Supplementum 360:90-95.

Ruggles D, Bharadwaj H, Shinn-Cunningham BG (2011) Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. Proc Natl Acad Sci U S A 108:15516-15521.

Ruggles D, Bharadwaj H, Shinn-Cunningham BG (2012) Why middle-aged listeners have trouble hearing in everyday settings. Current Biology 22:1417-1422.

Samelli AG, Matas CG, Carvallo RM, Gomes RF, de Beija CS, Magliaro FC, Rabelo CM (2012) Audiological and electrophysiological assessment of professional pop/rock musicians. Noise Health 14:6-12.

Smoorenburg, G. F. (1992). Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram. The journal of the acoustical society of America, 91(1), 421-437.

Spankovich C, Griffiths SK, Lobarinas E, Morgenstein KE, de la Calle S, Ledon V, Guercio D, Le Prell CG (2014) Temporary threshold shift after impulse-noise during video game play: Laboratory data. Int J Audiol 53 Suppl 2:S53-65.

Speaks C, Trine TD, Crain TR, Niccum N (1994) A Revised Speech Intelligibility Rating (RSIR) test: listeners with normal hearing. Otolaryngol Head Neck Surg 110:75-83.

Wang Y, Ren C (2012) Effects of repeated "benign" noise exposures in young CBA mice: shedding light on age-related hearing loss. J Assoc Res Otolaryngol 13:505-515.

Watson, C. S., Kidd, G. R., Miller, J. D., Smits, C., & Humes, L. E. (2012). Telephone screening tests for functionally impaired hearing: Current use in seven countries and development of a US version. *Journal of the American Academy of Audiology*, 23(10), 757-767.

Wilson RH (2011) Clinical experience with the words-in-noise test on 3430 veterans: comparisons with pure-tone thresholds and word recognition in quiet. J Am Acad Audiol 22:405-423.

Wilson RH, McArdle RA, Smith SL (2007) An Evaluation of the BKB-SIN, HINT, QuickSIN, and WIN Materials on Listeners With Normal Hearing and Listeners With Hearing Loss. J Speech Lang Hear Res 50:844-856.

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