A Meta-Analysis of the Efficacy of Music Therapy for Premature Infants

Jayne M. Standley, PhD, MT-BC

This meta-analysis on music research with premature infants in neonatal intensive care units (NICU) showed an overall large, significant, consistent effect size of almost a standard deviation (d = .83) (Cohen, 1998). Effects were not mediated by infants’ gestational age at the time of study, birthweight, or type of music delivery nor by physiologic, behavioral, or developmental measures of benefit. The homogeneity of findings suggests that music has statistically significant and clinically important benefits for premature infants in the NICU. The unique acoustic properties that differentiate music from all other sounds are discussed and clinical implications for research-based music therapy procedures cited.

Copyright 2002, Elsevier Science (USA). All rights reserved.

META-ANALYSIS IS A QUANTITATIVE synthesis of research data through formal statistical techniques, resulting in findings that are more objective than a traditional review of the literature. More definitive conclusions about the efficacy of specific clinical procedures can then be derived (Mann, 1990). The purpose of this article was to conduct a meta-analysis on the effects of music with premature infants and to determine if sufficient evidence exists to include music therapy in the clinical protocols of the neonatal intensive care units (NICU).

The literature contains a number of studies that have dealt with music’s effects on premature and low birthweight infants. Chapman (1975) played recorded lullabies in the incubator and found a 16% reduction in total time to reach weight criterion for discharge for premature infants. Similar studies have shown multiple benefits from short periods of continuously playing music in the incubator. Infants remain calmer and demonstrate more stable physiologic measures (Lorch, Lorch, Diefendorf, & Earl, 1994), have higher oxygen saturation levels (Cassidy & Standley, 1995; Collins & Kuck, 1991; Standley & Moore, 1995), gain weight faster (Malloy, 1979), and require a shorter time to reach discharge criteria (Caine, 1991). Additionally, the infants in Caine’s study were rated by their mothers to be calmer than those without music intervention at 6 months postdischarge follow-up (Standley, 1991).

Live singing with multimodal stimulation has been effective in enhancing developmental gains (Standley, 1998) and mother-infant bonding (Whipple, 2000). Contingent music increases non-nutritive sucking (Standley, 2000) and improves subsequent feeding ability (Standley, 1999). Finally, there is some evidence that music listening may reduce distress after frequent suctioning of premature infants requiring continuous ventilatory support (Burke, Walsh, Oehler, & Gingras, 1995).

MUSIC VERSUS NOISE IN THE NICU

Intentional sounds such as music are chosen for their potential for soothing, learning, and neurologic development. Ambient sounds are incidental, not chosen. They exist in the environment without controls for volume, duration, location, or cause/effect relationships, and can have unfortunate consequences: fatigue, stress, hyperalerting responses, startle, among others (Maschle, Rupp, & Hecht, 2000). Such ambient sound is especially detrimental in hospital settings (Cabrera & Lee, 2000; Novaes, Aronovich, Ferraz, & Knobel, 1997).

The medical community has concerns about the effects of ambient auditory stimuli on premature infants in the NICU, and guidelines have been drafted recommending greatly reduced decibel (dB) levels. These guidelines for sound stimuli do

From the Center for Music Research, Florida State University, Tallahassee, FL.
Address correspondence and reprint requests to Dr. Jayne M. Standley, School of Music, Florida State University, Tallahassee, FL 32306-1180. E-mail: jayne.standley@cmr.fsu.edu.
Copyright 2002, Elsevier Science (USA). All rights reserved.
0882-5963/02/1702-0005/$35.00/0
not differentiate ambient sound and noise from music nor do they incorporate known audiologic information regarding presentation of auditory stimuli to infants (Cassidy & Ditty, 1998). Basic acoustics demonstrate that music is different from all other sound (Farnsworth, 1958; Wagner, 1994).

Acoustic Properties of Music versus Noise

Sound is caused by vibration. Periodic, more complex vibrations, such as musical instruments and the human voice, produce harmonics perceived as pleasant, preferred auditory stimuli. Noise is a periodic vibration that results in irregular frequencies with inconsistencies of tension, stress, and configuration. Thereby, noise produces fatigue and stress in the listener. Ambient noise is the totality of the noises in one’s environment that is present but not chosen (Wagner, 1994).

Aural perception requires the translation of vibrations and is learned or developed over time. Music is an auditory stimulus with many cognitive elements such as melody, rhythm, harmony, timbre, form, style, and expressive characteristics that are processed simultaneously or in sequence. These cognitive elements are organized according to established rules of music within each culture. Repeated listening processes and identifies the organization and even allows the development of aural expectancies. It is theorized that this neurologic processing during music listening is a highly pleasurable activity until fatigue sets in. Acoustically, music is more pleasant, soothing, and interesting than noise; uses highly preferred frequencies and harmonics; and promotes neurologic organization (Wagner, 1994).

Therefore, music is often played in the NICU at slightly higher decibel levels than the ambient sound to mask aversive, stressful noise (Cassidy & Ditty, 1998). Unfortunately, the human voice frequencies are also masked by those of the incubator and other ambient noise; therefore, NICU infants are missing auditory processing opportunities, which may explain some of the later special education problems with communication disorders (Glass, 1994).

A descriptive literature review indicates that music would be beneficial in the clinical treatment of premature infants. However, comparison of music’s effects across all studies and generalization of diverse results to clinical treatment would be strengthened with systematic meta-analysis.

METHOD

Study Inclusion

Criteria for inclusion in this meta-analysis were: (1) experimental studies using group or individual subject designs; (2) subjects who were premature and low birthweight infants receiving treatment in a NICU; (3) music included as a separate, independent variable; and (4) reports in the English language of design, procedures, and results amenable to replicated data analysis.

The procedures followed the four basic steps of a meta-analysis: (1) a complete literature search was conducted to find all possible members of the defined population of studies whether in published or unpublished sources; (2) the characteristics and findings of the collected studies were identified, described, and coded; (3) these codes were independently reviewed for reliability and agreement, and then (4) each study’s results were statistically analyzed and converted to computed effect sizes using meta-analysis software (Johnson, 1989).

The identification process included exhaustive searches of the Journal of Music Therapy (1964-1999), Dissertation Abstracts International (1950-1999), and MEDLINE (1983-1999). Keywords for searches of electronic indexes included neonates, premature infants, NICU, music, and auditory stimulation. The reference lists of all relevant articles located in the electronic search were also reviewed.

Study Descriptions

Ten studies met criteria for inclusion in the meta-analysis and are summarized here.

Caine (1991) studied 52 premature infants. Twenty-six were exposed to 30 minutes of recorded lullabies alternated with 30 minutes of ambient auditory stimulation 3 times per day throughout their stay in the NICU and 26 were exposed only to ambient auditory stimulation. As compared to the control infants, those exposed to music had reduced initial weight loss, increased daily weight gain, increased formula intake, and significantly reduced length of hospitalization.

Cassidy and Standley (1995) played recorded lullaby music for 20 oxygenated infants of 24 to 30 weeks’ gestational age for 3 days during the infants’ first week of life. Oxygen saturation levels, heart rate, respiratory rate, and number of apnea/bradycardia episodes were recorded once per minute throughout nine 4-minute segments alternating music with 4-minute segments of ambient noise. Results indicated no contraindications for
music in the first week of life for premature infants at this early age and showed positive effects on all physiologic measures. Apnea/bradycardia episodes were not increased by music treatment.

Coleman, Pratt, Stoddard, Gerstmann, and Abel (1997) studied 66 infants. Half were exposed for 4 days to three randomly ordered 20-minute segments of auditory stimulation consisting of males and females singing lullabies versus males and females speaking lullabies versus ambient NICU sounds. The infants exposed to music left the NICU 3 days earlier than the control group and showed significantly higher caloric intake and weight gain. Infants responded equally to male and female voices.

Collins and Kuck (1991) combined recorded uterine sounds with female lullaby singing for 17 infants. When these infants were in an agitated state, they were observed with and without the music intervention. Oxygen saturation and behavioral state improved significantly under the music condition.

Flowers, McCain, and Hilker (1999) played 20 minutes of soft-rock ballads from the radio versus 20 minutes of music combined with uterine sounds for nine African-American premature infants serving as their own controls. Results showed significant differences for the womb music for respiratory rate, oxygen saturation, and time spent sleeping during auditory stimulation. No significant differences were found for heart rate or number of behavioral state changes.

Moore, Gladstone, and Standley (1994) compared mothers reading poetry versus lullabies sung by females versus white noise versus uterine sounds versus silence. All auditory stimuli were in 10-minute segments randomly ordered and all five were played for each of 10 infants. All auditory tapes resulted in significantly higher oxygen saturation levels compared to the silence condition. There were no significant differences in oxygen saturation among the four sound stimuli.

Standley and Moore (1995) studied 20 oxygenated low birthweight infants exposed to 20 minutes of auditory stimuli for 3 consecutive days. Ten listened to lullabies sung by females and 10 listened to their mothers’ voices reading. On day 1 the music showed significantly higher oxygen saturation levels than the mothers’ voices but by days 2 and 3 the mothers’ voices were achieving results similar to the music. Infants hearing music had significantly fewer occurrences of Oximeter alarms than did infants listening to their mothers’ voices.

Standley (1998) combined live lullaby singing with multimodal stimulation for 40 premature infants who were at least 32 gestational weeks at the time of the study. Half received the stimulation and half were in the control condition. The infants receiving stimulation left the hospital significantly sooner than the infants in the control group with significant gender differences. Music females left an average of 11.9 days earlier than females in the control group and music males left 1.5 days earlier than control males. Also, music infants gained significantly more weight per day than control infants.

Standley (2000) developed a pacifier attached to an air pressure transducer and timer such that a change in air pressure induced by a nonnutritive suck activated a tape recorder that played 10 seconds of a lullaby sung by a female. Twelve infants were exposed to an ABAB design alternating baseline (A) (non-nutritive sucking [NNS] with no music reinforcement) with the contingent music condition (B) followed by a full reversal (AB). Results showed that infants sucked significantly more when reinforced with the 10 seconds of music than during the no-music sucking condition. Additionally, the infants (average gestation age of 35.2 weeks) required an average of 2.5 minutes to learn to suck so that the music was continuously playing.

Standley (1999) used the pacifier described in the previous study to compare feeding rates for 32 premature infants at approximately 36 weeks’ gestational age. All infants were attempting two bottle feedings per day and were referred by the physical therapist for failure to independently feed. Rates of feeding were computed for the morning and evening bottle feeding attempts within a 24-hour period. Half of the infants received a 15-minute opportunity for music-reinforced nonnutritive sucking approximately 30 to 60 minutes before the evening bottle-feeding attempt. Results demonstrated significantly increased feeding rates after music-reinforced nonnutritive sucking.

**Data Extraction**

The value of each dependent variable reported in the selected studies was converted to the estimated effect size, d (Cohen, 1988). Table 1 identifies each study, its sample size, dependent variables, study statistics converted to effect size, and the resulting d and Pearson r statistics. Table 2 shows each study’s date of publication, publication status, independent variable, type of music with mode of delivery, reported decibel levels, mean group adjusted gestational age of subjects at the time of the study, mean group birthweight, and type of data.
analysis, either between (B) or within (W) subjects.

**RESULTS**

All effect sizes were in a positive direction for the effects of music and ranged from .4915 to 1.9528. Analysis of the effect sizes revealed an overall $d = .83$, the difference between the music and nonmusic results in standard deviation units (Table 3). Because the 95% confidence interval did not include 0, this effect size is considered statistically significant and indicates that music gener-

### Table 1. Effect Size Analysis by Study

<table>
<thead>
<tr>
<th>Study</th>
<th>Study n</th>
<th>Dependent Variable</th>
<th>Study Statistic</th>
<th>Effect Size $d$</th>
<th>Effect Size Pearson $r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caine (1991)</td>
<td>52</td>
<td>Days in hospital</td>
<td>$t = 1.847$</td>
<td>.5045</td>
<td>.2481</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weight gain</td>
<td>$t = 3.066$</td>
<td>8375</td>
<td>.3913</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Behavior state</td>
<td>$t = 2.666$</td>
<td>7283</td>
<td>.3468</td>
</tr>
<tr>
<td>Cassidy &amp; Standley (1995)</td>
<td>20</td>
<td>Oxygen saturation</td>
<td>$F = 7.70$</td>
<td>1.1885</td>
<td>.5272</td>
</tr>
<tr>
<td>Coleman, Pratt, Stoddard, Gerstmann, &amp; Abel (1997)</td>
<td>66</td>
<td>Heart rate</td>
<td>$F = 14.27$</td>
<td>9190</td>
<td>.4216</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Oxygen saturation</td>
<td>$F = 12.60$</td>
<td>8636</td>
<td>.4004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Behavior state</td>
<td>$F = 64.43$</td>
<td>1.9528</td>
<td>.7028</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Days in hospital</td>
<td>$p &lt; .05$</td>
<td>.4915</td>
<td>.2413</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weight gain</td>
<td>$p &lt; .05$</td>
<td>.4915</td>
<td>.2413</td>
</tr>
<tr>
<td>Collins &amp; Kuck (1991)</td>
<td>17</td>
<td>Oxygen saturation</td>
<td>$p &lt; .05$</td>
<td>.6971</td>
<td>.3362</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Behavior state</td>
<td>$t = 3.75$</td>
<td>1.2559</td>
<td>.5409</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Heart rate</td>
<td>$t = 1.36$</td>
<td>4555</td>
<td>.2271</td>
</tr>
<tr>
<td>Flowers, McCain, &amp; Hilker (1999)</td>
<td>9</td>
<td>Oxygen saturation</td>
<td>$F = 5.74$</td>
<td>1.0503</td>
<td>.5026</td>
</tr>
<tr>
<td>Moore, Gladstone, &amp; Standley (1994)</td>
<td>22</td>
<td>Oxygen saturation</td>
<td>$F = 4.20$</td>
<td>8809</td>
<td>.4383</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weight gain</td>
<td>$F = 6.833$</td>
<td>8102</td>
<td>.3820</td>
</tr>
<tr>
<td>Standley (2000)</td>
<td>12</td>
<td>Nonnutritive sucking rate</td>
<td>$F = 7.460$</td>
<td>.7334</td>
<td>.3668</td>
</tr>
<tr>
<td>Standley (1999)</td>
<td>32</td>
<td>Feeding rate</td>
<td>$t = 2.252$</td>
<td>8726</td>
<td>.4085</td>
</tr>
<tr>
<td>Standley &amp; Moore (1995)</td>
<td>20</td>
<td>Oxygen saturation</td>
<td>$t = 2.40$</td>
<td>1.0280</td>
<td>.4729</td>
</tr>
</tbody>
</table>

### Table 2. Study Characteristics and Codes

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Published or in Press</th>
<th>Independent Variable</th>
<th>Music Type/Delivery</th>
<th>dB Level</th>
<th>Adjusted Gestational Age at Study</th>
<th>Birthweight in Grams</th>
<th>Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caine (1991)</td>
<td>1991</td>
<td>Yes</td>
<td>Lullabies vs. RAS</td>
<td>Recorded/free field</td>
<td>70-80</td>
<td>NA</td>
<td>1,678</td>
<td>B</td>
</tr>
<tr>
<td>Cassidy &amp; Standley (1995)</td>
<td>1998</td>
<td>Yes</td>
<td>Lullabies vs. RAS</td>
<td>Recorded/free cups</td>
<td>80</td>
<td>27 weeks</td>
<td>&lt;1,361</td>
<td>B</td>
</tr>
<tr>
<td>Coleman et al. (1997)</td>
<td>1997</td>
<td>Yes</td>
<td>Lullabies vs. RAS</td>
<td>Recorded/free field</td>
<td>65-75</td>
<td>29.5 weeks</td>
<td>NA*</td>
<td>B</td>
</tr>
<tr>
<td>Collins &amp; Kuck (1996)</td>
<td>1996</td>
<td>Yes</td>
<td>Lullabies with heart beats</td>
<td>Recorded/free field</td>
<td>80</td>
<td>30 weeks</td>
<td>1,751</td>
<td>W</td>
</tr>
<tr>
<td>Flowers et al. (1999)</td>
<td>1999</td>
<td>No</td>
<td>1970s ballads and lullabies with heart beats</td>
<td>Recorded/free field</td>
<td>55</td>
<td>28 weeks</td>
<td>932</td>
<td>W</td>
</tr>
<tr>
<td>Moore et al. (1997)</td>
<td>1997</td>
<td>No</td>
<td>Lullabies with white noise</td>
<td>Recorded/earphones</td>
<td>75-80</td>
<td>31 weeks</td>
<td>NA</td>
<td>W</td>
</tr>
<tr>
<td>Standley (1998)</td>
<td>1998</td>
<td>Yes</td>
<td>Lullabies with cephalocaudal massage vs. RAS</td>
<td>Live singing</td>
<td>Quiet</td>
<td>30.6 weeks</td>
<td>1,298</td>
<td>B</td>
</tr>
<tr>
<td>Standley (2000)</td>
<td>2000</td>
<td>Yes</td>
<td>Lullabies contingent upon pacifier suck</td>
<td>Recorded/free field</td>
<td>65-70</td>
<td>35.5 weeks</td>
<td>1,747</td>
<td>W</td>
</tr>
<tr>
<td>Standley (1999)</td>
<td>1999</td>
<td>No</td>
<td>Lullabies contingent upon pacifier suck</td>
<td>Recorded/free field</td>
<td>65-70</td>
<td>36.1 weeks</td>
<td>1,495</td>
<td>B</td>
</tr>
</tbody>
</table>

Abbreviations: B, Between subjects; W, Within subjects.

Note. *It is unclear whether dB levels were determined with scale A or scale C measurements.

*Mean group age.

*Mean group weight.

*Routine auditory stimulation.

*Not applicable.
ally has a positive and significant impact in the NICU. The Q value was not significant \((p = .1752)\), which means that the effect sizes of music studies in the NICU were consistent and adequately explained by the single, mean effect size.

Results demonstrate that music therapy procedures provide significant benefits to premature infants that are consistent across the following diverse variables: the adjusted gestational age of the infant at the time of music intervention; the decibel level of the music within a range of 55 to 80 dB; the mode of delivery (free field/earphones); and the birthweight of the infant. Additionally, results are consistent across all variables measured: observed behavioral state, heart rate, respiration rate, oxygen saturation level, weight gain, days in hospital, feeding rate, and nonnutritive sucking rate.

Table 4 is a summary of clinical programs and guidelines for use of music in the NICU that were garnered from these research results. Results of this meta-analysis justify the incorporation of these music therapy programs into the standard of care for premature infants.

<table>
<thead>
<tr>
<th>Clinical Implications for Health Professionals</th>
</tr>
</thead>
</table>

Analytic results of the research in music with premature infants provide multiple implications for their health care. Three basic procedures are documented as beneficial. First is to play recorded, carefully selected lullabies (see Table 4) to pacify infants, to increase oxygen saturation, to reduce stress, or to increase language stimulation. Such music applications could begin as early as 28 weeks’ corrected gestational age unless contraindicated by the individual child’s condition. Contraindications would include observed hyperresponsiveness to music specifically (as opposed to other sounds and noises) and known inability to hear. The music should be provided in short intervals of 20 to 30 minutes at critical periods throughout the 24-hour day. Critical periods might include the beginning of sleep and quiet times and the periods immediately after stressful medical and daily care procedures, such as feeding, suctioning, or eye exams.

Second is to sing lullabies as a mechanism to sustain homeostasis during multimodal stimulation. Such stimulation could begin at approximately 32 weeks’ gestational age for enhancing neurologic development and promoting tolerance for increasing stimulation. Contraindications would include severe hyperresponsiveness to touch and medical conditions preventing touch.

Third is to use music as reinforcement for nonnutritive sucking to increase sucking ability and duration (Standley, 2000) and to increase feeding

### Table 3. Overall Results of Meta-Analysis of Music with Premature Infants

<table>
<thead>
<tr>
<th>Effect</th>
<th>Effect</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size d</td>
<td>95% CI</td>
<td>Size r</td>
</tr>
<tr>
<td>Overall Results</td>
<td>0.268</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Note. \(Q(19) = 24.57; p = .1752\). Total \(N = 780\) subjects; mean \(N/\text{study} = 39\) subjects.
ability (Standley, 1999). This procedure is documented in the research literature as being effective as early as 30 to 32 weeks’ gestational age. Contraindications would include medical conditions precluding pacifier sucking, such as use of the ventilator.

Music therapy can provide other benefits in the care of premature infants, including promoting bonding with family members; counseling parents; educating parents about infant care, pacification, and developmental stimulation; and assisting with the infant’s transition to the home environment. These procedures, which combine counseling skills with knowledge of the use of music to accomplish therapeutic and educational objectives, would require the professional services of a board-certified music therapist.

NICU health care professionals should be cautioned about other uses of music undocumented in the research literature as beneficial to premature infants receiving intensive care. These include live instrumental performances in the NICU environment, radio stations tuned to selections preferred by the medical staff rather than selected for the benefits of the infants, and classical music selections widely acclaimed to enhance intelligence. These must be subjected to controlled research procedures before benefits can be claimed. At this point, their clinical importance is unknown.

**DISCUSSION**

Further research with music and premature infants is definitely warranted. This meta-analysis had a mean of 39 subjects per study, which is somewhat low. A variety of research studies with a larger subject pool is necessary to clarify the efficacy of specific procedures for music exposure, especially regarding prime gestational age, length of exposure, and decibel levels. Additionally, a larger pool of studies would allow for analysis for a particular dependent variable. Long-term follow-up of infants receiving music in the NICU is definitely needed and, as yet, has not been documented in the literature.

This meta-analysis provided strong and definitive results; however, its importance is, to some extent, constrained by the limitations of the studies reviewed. The majority of studies did not include diagnostic and condition-related factors of subjects beyond gestational age at birth and birthweight. Therefore, it is difficult to judge equality of subject groups with regard to severity of medical status. A few studies involved a single presentation of auditory stimuli. Such results may have been due more to the presentation of a novel stimulus than to music per se. Studies with multiple intervention opportunities and more carefully controlled subject groupings would seem warranted.

In general, this study provides support for the clinical use of a variety of music therapy techniques with premature infants. Perhaps one of the most important is, as yet, undocumented in the research literature: that lullaby selections for premature infants convey the human voice and may provide much-needed language stimulation with important long-term consequences for future learning. Music alone or combined with the human voice would seem to be a valuable resource for enhancing developmental goals in the NICU and functioning to reduce stress, to provide developmental stimulation during a critical period of growth, to promote bonding with parents, or to facilitate neurologic, communication, and social development.

**REFERENCES**


Glass, P. (1994). The vulnerable neonate and the neonatal intensive care environment. In G.B. Avery, M.A. Fletcher, & M.G. MacDonald (Eds.). *Neonatology: Pathophysiology and...*


