Benefits of Massage Therapy for Infants With Symptoms of Gastroesophageal Reflux Disease

Madalynn Neu, PhD, RN¹, Zhaoxing Pan, PhD², Rachel Workman, MD³, Cassandra Marcheggiani-Howard, LSW¹, Glenn Furuta, PhD, MD²,⁴, and Mark L. Laudenslager, PhD²

Abstract

Objectives: This randomized controlled pilot trial was conducted to evaluate the clinical efficacy of massage therapy (MT) for relief of symptoms of gastroesophageal reflux disease (GERD). The hypothesis was that, when compared to infants who received nonmassage therapy, infants who received MT would display fewer GERD symptoms, greater weight gain, greater amount of sleep, lower cortisol levels before and after treatment, and lower daily (area under the curve [AUC]) cortisol secretion.

Methods: Participants were 36 infants born at term, 4–10 weeks of age at enrollment, healthy except for a diagnosis of GERD by their pediatrician, and with a score of at least 16 on the Infant Gastroesophageal Reflux Questionnaire–Revised (I-GERQ-R). Infants were randomized to receive either MT or a nonmassage sham treatment in their homes for 30 min twice a week for 6 weeks. Data collectors and parents were blind to study condition. Results: GERD symptoms decreased in both groups and weight increased. Pretreatment salivary cortisol levels decreased significantly over time in the massage group while increasing in the nonmassage group. Daily cortisol level also decreased in the massage group and increased in the nonmassage group, but the difference was not significant. Conclusions: MT administered by a professional therapist did not affect symptoms of GERD differently than a sham treatment but did decrease infant stress as measured by cortisol. Research focusing on stress reduction in infants with GERD and multimodal treatments addressing GERD symptoms may yield the most effective treatment.

Keywords
infant, gastroesophageal reflux, GERD, massage therapy, randomized controlled trial

Almost 70% of infants under 6 months of age experience gastroesophageal reflux, or GER (Nelson, Chen, Syniar, & Christoffel, 1997). GER occurs when the lower esophageal sphincter relaxes transiently, allowing gastric contents to flow back into the esophagus. Immaturity of the stomach, a high liquid diet, and esophageal immaturity contribute to GER in infants (Vandenplas et al., 2009). Approximately 25% of infants experience GER complications, called gastroesophageal reflux disease (GERD).

One manifestation of GERD in infants is irritability (e.g., crying, fussing, and back arching). In a study of 185 infants diagnosed with GERD, 70% of mothers reported infant irritability (Kleinman et al., 2006). Typical healthy infants cry for 2–2½ hr a day at 6–8 weeks of age and 1 hr a day by 12 weeks (Walker & Menahem, 1994). In contrast, investigators found that infants less than 9 months of age with GERD cried 4.2–4.5 hr per day before intervention and 2.5–3 hr after intervention (Jordan, Heine, Meehan, Catto-Smith, & Lubitz, 2006; Orenstein et al., 2003). Irritability may occur because of discomfort from acid reflux, but research has shown that crying and fussing are temporally associated with bouts of both acid and nonacid reflux (Condino et al., 2006). Another GERD manifestation in infants is feeding difficulty (e.g., coughing, choking, feeding refusal, and gagging). Such feeding problems and/or frequent regurgitation can result in insufficient caloric intake and inadequate weight gain (Tolia, Wuethr, & Thomas, 2003). Irritability and frequent regurgitation can interfere with sleep. Infants with GERD experience more episodes of acid and nonacid reflux that produce awakenings and delayed onset of sleep than other infants (Machado et al., 2013).

¹University of Colorado Anschutz Medical Campus, College of Nursing, Denver, CO, USA
²Department of Pediatrics, University of Colorado, School of Medicine, Anschutz Medical Campus, Denver, CO, USA
³Pediatrics, Denver, Colorado
⁴Gastrointestinal Eosinophilic Diseases Program, Children’s Hospital Colorado, Aurora Colorado

Corresponding Author:
Madalynn Neu, PhD, RN, University of Colorado College of Nursing, 13120 E. 19th Ave. Box C288-19, Aurora, CO, USA.
Email: madalynn.neu@ucdenver.edu
A heightened sensation of reflux in the esophagus can increase infant stress (Hamilton & Zeltzer, 1994) regardless of acid level of the reflux. Feeding discomfort/unpleasantness, feeding refusal resulting in hunger, and awareness of unmet basic needs are other potential reasons for stress in infants with GERD. Frequent challenges may overwhelm infants resulting in frequent or continuous stress (Hamilton & Zeltzer, 1994).

Cortisol, the primary biomarker of hypothalamic–pituitary–adrenocortical (HPA) axis activity, is the end product of HPA-system activation in response to stress or emotional arousal. The negative feedback mechanism initiated by an elevation of cortisol levels prevents long-term elevation (McEwen, 1997); however, chronic HPA-system activation interferes with negative feedback and thus with cortisol-level regulation.Infants show increases in cortisol levels in response to stress (Elverson, Wilson, Hertzog, & French, 2012; Towe-Goodman, Stifter, Mills-Koonce, & Granger, 2012) and, because of immaturity of HPA function (Larson, White, Cochran, Donzella, & Gunnar, 1998), they are especially at risk for subsequent dysregulation of the HPA system when exposed to frequent or continual stress.

Infant irritability is a primary reason that mothers consult a pediatrician (Kleinman et al., 2006). Irritability plus regurgitation typically lead pediatricians to diagnose GERD (Diaz et al., 2007) and then treat it either nonpharmacologically or pharmacologically. Nonpharmacologic treatments include upright positioning, which may worsen regurgitation by increasing pressure on the stomach, and feedings thickened with rice cereal, which is not practical for breast-fed infants (Orenstein, Whittington, & Orenstein, 1983). Antireflux medications, such as histamine H₂ receptor antagonists (H₂RA) or more potent proton-pump inhibitors (PPIs), often are prescribed to decrease esophageal acid exposure (Diaz et al., 2007). While studies on safety in infants for these medications are scant, in one investigation 32% of infants taking an H₂RA exhibited adverse effects such as agitation, head rubbing, somnolence, diarrhea, and vomiting (Orenstein et al., 2003). In a blinded, placebo-controlled study of infants with GERD, symptom relief from lansoprazole treatment was no different from placebo (Orenstein, Hassall, Furmaga-Jablonska, Atkinson, & Raanant, 2009). Furthermore, the Advisory Committee of the Food and Drug Administration no longer recommends PPIs for treatment of symptoms of GERD in healthy infants without documented evidence of acid reflux (Chen et al., 2012). Thus, research has shown that neither standard nonpharmacologic nor pharmacologic treatment is more effective in relieving infant irritability (the mother’s reason for seeking care) than is placebo.

An alternative treatment that has potential to improve mood and increase relaxation/sleep in infants is massage therapy (MT), the mechanical manipulation of body tissues with rhythmic pressure and stroking (Field, Diego, & Hernandez-Reif, 2010). In one study, stress behaviors (including crying) decreased and weight gain increased in newborns who received 15 min of massage by a professional therapist daily for 10 days (Scafidi & Field, 1996). In another, massage administered twice a week for 6 weeks promoted more positive mood, weight gain, and lower cortisol levels in infants 4–12 weeks of age (Field, Grizzle, Scafidi, Abrams, & Richardson, 1996). Crying decreased (vs. control group) in studies where mothers administered massage daily for 1 week (Arikan, Alp, Gozum, Orbak, & Cifci, 2008) and 5 weeks (Çetinkaya & Buşbakakk, 2012). And in a study in which mothers massaged their newborn infants daily for 2 weeks, infants in the intervention group displayed better circadian sleep patterns (as assessed by actimetry) at 8 and 12 weeks of age than those in a control group (Ferber, Laudon, Kuint, Weller, & Zisapel, 2002). Interestingly, when investigators measured salivary cortisol in 14 healthy newborn infants before and after only one massage session, they found that levels increased after the session (White-Traut, Schwertz, McFarlin, & Kogan, 2009), which may have been due to either the novelty of the treatment or the stimulation associated with the manipulation.

While researchers have not explored the effects of massage specifically among infants who have GERD symptoms, the literature suggests that massage has the potential to relieve infant GERD symptoms and enhances outcomes such as mood, sleep, weight gain, and decreased stress levels (Arikan et al., 2008; Ferber et al., 2002; Field et al., 1996; Scafidi & Field, 1996). The objective of the present study was to evaluate the clinical efficacy of MT in relieving GERD symptoms in infants. We hypothesized that, when compared to infants with GERD who did not receive massage, those with GERD who received MT would display (a) a greater decrease in GERD symptoms (Hypothesis 1), (b) a greater weight gain (Hypothesis 2), (c) a greater increase in amount of sleep (Hypothesis 3), (d) a greater decrease in cortisol levels measured before treatments (Hypothesis 4), and (e) lower overall cortisol concentrations (Hypothesis 5). Our second objective was to evaluate the feasibility of testing massage as a treatment for GERD among infants (i.e., procedure fidelity, blinding, dose response, maternal acceptance, accrual, and attrition).

Materials and Methods
We conducted a randomized controlled trial feasibility study to compare the effects of MT to a nonmassage therapy (NMT) sham treatment on the primary outcome, GERD symptoms, and three secondary outcomes, weight, sleep, and salivary cortisol.

Sampling Plan and Recruitment
We planned our target sample of 36 (18 per group) to ensure 80% power at 5% significance to detect the between-group difference of an effect size of 1.0, the common standard deviation (SD) in analysis of covariance (ANCOVA) adjusted for baseline value. Pediatric-care providers in 11 pediatric practices in the metropolitan area of a mid-sized city in the western United States referred mothers of eligible infants to the study. We also advertised through university campus e-mail. Infants were included in this study if they (a) were born at 38–42 weeks gestational age, (b) were 6–10 weeks of age at enrollment, (c) had been diagnosed with GERD by their pediatric provider, and (d) had scored at least 16 on the Infant Gastroesophageal Reflux Questionnaire–Revised (I-GERQ-R). Mothers were
Neu et al. 389

(a) English speaking and (b) at least 18 years of age. Infants were not eligible if they (a) had undergone major surgery, (b) had been diagnosed with a chronic illness other than GERD, (c) had a congenital anomaly, (d) had experienced bilious or projectile vomiting or bloody stool, (e) were hospitalized for a condition other than GERD, or (f) were taking steroidal medications. The Colorado Multiple Institution Review Board approved the research, and a research assistant (RA) obtained written informed consent at first data collection. The project coordinator used a computerized table of random numbers to determine group assignment after consent was obtained. Opaque envelopes concealed allocation. Parents and RAs who collected and entered data were blinded to group assignment of the infants.

Treatment Procedures

Two certified therapists, experienced in infant massage, performed massage. Two therapists performed NMT, a graduate nursing student experienced in infant care and a pediatric physical therapist. Infants were assigned to one of these therapists, who performed all treatments on the infant. Within treatment groups, we rotated infant assignments among therapists, with each therapist typically being assigned to every other infant enrolled. The principal investigator (PI) instructed therapists in their respective protocols and practiced with them using a doll until they were 100% consistent with the protocol. Treatment sessions were videotaped, and the project coordinator reviewed all videos for adherence to the protocol. Two RAs watched 25% of the videos independently and recorded adherence to each item of the protocol using yes/no responses. Percentage adherence was calculated independently by the coders and by the therapists.

Therapists scheduled treatments in late afternoon or evening twice a week for 6 weeks, 90 min after the most recent feeding to limit the potential of reflux during the session. Treatment took place in a quiet room in the infant’s home, usually in the infant or parent’s bedroom. No family members were in the room during treatment, but a nursing student accompanied the therapist to assist with videotaping. After the 6-week data collection, we revealed the treatment-group membership to mothers and offered a massage demonstration.

During both massage and nonmassage therapies, each infant could have a pacifier. Therapists used a warming pad on low heat covered with a blanket if the room felt cool. They responded to smiles, eye contact, and infant vocalizations but avoided initiating verbalization. If the infant cried incoherably for 3–4 min, became apneic/dusky, or had several bouts of reflux, the therapist terminated treatment and rescheduled for as soon as possible during the same week.

The massage protocol was an extended version of the 15-min protocol detailed by Field, Grizzle, Scafidi, Abrams, and Richardson (1996). Therapists used moderate hand pressure (Field et al., 2010), administering each of the six steps in the session (face and head, chest, abdomen, legs and feet, arms and hands, and back) for 5 min, for a total duration of 30 min. The extension of our protocol to 30 min was informed by the usual practices of the therapists. The nonmassage sham protocol also lasted 30 min. During the first 10 min, the therapist placed one hand over the other, administering light, consistent pressure for 1 min each on the infant’s forehead, upper arms (one at a time), chest, abdomen, each thigh, each shoulder, and back. Then the therapist held the infant vertically on her shoulder for the remaining 20 min. The sham treatment was similar to the rocking described by Field et al. (1996) and to the touching and holding mothers typically perform.

Measures

I-GERQ-R. We used the parent report on the I-GERQ-R to measure our primary outcome of GERD symptoms. The instrument consists of 12 questions addressing GERD symptoms during the week prior to I-GERQ-R administration. Questions address (a) amount of daily crying or fussing (b) instances of regurgitation, arching back, refusal to feed, or stopped feeding, hiccups, and apnea. Response choices for most questions are 0 (never) to 4 (always). The possible score ranges from 0 to 42. Higher scores indicate a greater burden of symptoms. Validation was conducted in seven countries with 185 GERD-diagnosed infants and 93 control infants. Internal consistency and test–retest reliability were greater than 0.85. A cutoff score of 16 yielded sensitivity of 0.65 and specificity of 1.0 (Kleinman et al., 2006).

Weight. We obtained infant weight with a Detecto portable infant scale, model 8440 (Cardinal Scale; Warwick, RI). The scale is accurate to 0.2% ± 1 digit of readings over 200 g. We tared the scale to 0 with a blanket and clean diaper on it and then weighed the unclothed baby.

Actigraphy. We used the Actiwatch 2 actigraph system (Phillips Respironics; Bend, OR) to measure total amount of sleep. Polysomnography (PSG), gold standard for sleep measurement, must be conducted in a laboratory (Crowell et al., 1997), but the actiwatch is portable and records sleep during the course of regular daily activities. The instrument resembles a small watch and utilizes an accelerometer to monitor the frequency and intensity of movement. We used the autoactivity setting of the actiwatch to assess total amount of sleep because agreement with PSG for infants aged 2–3 months is greater than 89% for total sleep on that setting (So, Buckley, Adamson, & Home, 2005).

Salivary Cortisol Level. Saliva was collected using filter paper that has been validated for use with infants (Whatman grade 42 filter paper, 2.54 cm x 9.0 cm, GE Healthcare, Waukesha, WI; Neu Goldstein, Gao, & Laudenslager, 2007). Cortisol concentration was determined using a commercial expanded range high-sensitivity EIA kit (no. 1-3002/1-3012, Salimetrics) that detects cortisol levels in the range of 0.003–3.0 μg/dl with a low-end detection limit after extraction of approximately 0.015 μg/dl. Unknowns were determined from weighted regression analysis of the standard curve using commercial software (Revelation 3.2) for the ELISA plate reader (Dynex MRX). Laboratory controls were run on every plate for
determination of interassay and intraassay coefficients of variability, which were generally less than 5% and 9%, respectively.

Additional Measures. To more fully understand the infant’s response throughout the data collection and to address potential confounding variables, we asked mothers to complete daily diaries that addressed how often the baby displayed regurgitation, crying, and fussing. We used the data to track changes in GERD symptoms relative to timing of massage or nonmassage interventions. Once a week, mothers completed a set of questions relating to potentially confounding matters, such as use of antireflux medications, nonpharmaceutical treatments prescribed by the pediatric provider, or alternative treatments the mother may have tried.

Mothers were not screened for postpartum depression or increased anxiety prior to enrolling in the study. We did, however, measure maternal anxiety and depression as potential confounds because they have been associated with infant feeding problems (Karacetin, Demir, Erkan, Cokugras, & Somnez, 2011). To measure maternal anxiety, we used the State-Trait Anxiety Inventory (STAI), a 40-item questionnaire that assesses presence of current state anxiety and more stable trait anxiety. Well-established validity and reliability have been reported for various ages and ethnicities (Spielberger, 1983). To measure postpartum depression, we used the Edinburgh Postnatal Depression Scale (EPDS), a 10-item self-report scale (score range 0–30) assessing the common symptoms of postpartum depression. Validity, reliability, positive predictive value, and sensitivity and specificity with a cutoff point of 12 are well established (Murray & Carothers, 1990).

Procedures

Massage and nonmassage treatments were given twice a week for 6 weeks. Data were collected at baseline and after the second weekly massage or nonmassage treatment at 4 and 6 weeks. Mothers were compensated US$40 at each assessment. The schedule and procedures for collecting specific data were as follows.

Questionnaires. Mothers completed the demographic form, STAI, and EPDS at baseline and at 6 weeks. They completed the I-GERQ-R at baseline, Week 4, and Week 6. They answered open-ended questions about study acceptability at Weeks 4 and 6. The therapist collected daily and weekly diaries after the second session each week.

Activwatch. The RA placed the activwatch on infants at baseline and Weeks 4 and 6. For 3 days at each time point, infants wore the actigraph watch over a sock on the right leg just above the ankle. Mothers recorded the infants’ sleep/awake times, potential confounds (e.g., riding in the car and swaddling), and watch removal times (to bathe the infant) in a diary.

Weight. The RA weighed the infant at baseline and Weeks 4 and 6.

Salivary Cortisol. The PI trained the therapists to collect saliva samples. They collected one sample from the infant immediately before massage or nonmassage treatment at baseline and Weeks 4 and 6 using filter paper. Filter papers were contained in a booklet (one booklet for each data collection: baseline, 4 weeks, 6 weeks), as described by Laudenslager, Calderone, Phillips, Natvig, and Carlson (2013). Before collecting the saliva sample, the therapist gently wiped the infant’s mouth with a soft, dampened cloth to remove any residual formula. The therapist then spread the front and back covers of the booklet open so the filter could be placed on the infant’s tongue while remaining in the booklet. Infants sucked on the paper for approximately 30–60 s. After assuring that at least 1 inch of the paper was saturated, the therapist removed the paper, marked the exact time that the sample was obtained on the booklet, closed the booklet, and placed it in a plastic bag with holes cut into the bag to facilitate drying of the filter paper. Filters were stored at room temperature until assayed (Neu et al., 2007). At each data collection session, therapists also completed a short questionnaire that asked whether the infant was crying before collection, when the infant had most recently eaten, and what medications the infant had taken the day of collection.

Mothers collected saliva samples from their infants 3 times daily before feedings (infant morning awakening, 12 p.m., and 10 hr after awakening) on 3 consecutive days before the first treatment at baseline and after the last treatment at Week 6. The RA explained the procedure to mothers during baseline data collection, including instructions to collect the sample at least 1 hr after a feeding and wipe out the infant’s mouth to remove residual formula. RAs also gave mothers written instructions and told mothers to call them with any questions. Mothers received three filter booklets, each labeled with a tab indicating one of the three specific days of collection. Each booklet contained three filters, one for each time of daily collection. Mothers were told to record the date and exact time of each sample collection on the booklet and to store each booklet in a separate ventilated plastic bag. When they used this technique in a previous study, Laudenslager et al. (2013) reported an $r^2$ of .98 between the time recorded by the subject directly on the booklet and the time recorded by an electronic collection device for 286 observations. In the present study, mothers also completed a saliva survey each collection day, recording factors that could interfere with cortisol levels (e.g., infant medications, activity level, and time of last feeding).

Statistical Analysis

Sleep data were averaged over the 3-day collection period. If more than 50% of sleep data were missing, we did not use that day in the analysis. Daily cortisol data collected over 3 days were averaged to reduce variability, and area under the curve (AUC) with respect to the ground of these aggregated cortisol levels collected over a 10-hr period was calculated using the trapezoidal rule.

Statistical analyses were performed using SAS 9.2 software (SAS Institute, Inc., Cary, NC). We have presented descriptive
statistics in tables as either mean ± SD for continuous measures or percentages for categorical responses. Two-sample t-test or the χ² test was used to examine potential differences in baseline values and demographic variables between the two groups. We examined normality of outcome variables and found that cortisol levels were not normally distributed; hence, we applied logarithmic transformation to a natural base to the cortisol outcome before statistical analysis. ANCOVA adjusting for the baseline value was used to assess efficacy of the intervention on a particular outcome variable assessed at Week 6. Least square estimate was used to quantify the between-group difference. Last observation carry forward was used. Repeated measures analysis of variance with linear mixed-effects model was used to test the profile of the change in outcome variable at baseline, 4 weeks, and 6 weeks. AUC of daily cortisol could not be normalized with log transformation because of two extreme outliers. Wilcoxon rank-sum test and Hodges–Lehmann estimator were used to assess the between-group difference in the postintervention AUC change of daily cortisol. Significance level was set at α = .05.

Results

Sample

Figure 1 shows details of recruitment. Of the 70 mothers we recruited, 43 consented to participate. We randomized 21 to the massage and 22 to the nonmassage group, and 18 and 17 mothers in each group, respectively, completed the study. In addition, 1 mother in the nonmassage group withdrew after completing data collection at Week 4. Because we used intention-to-treat analysis, we used her 4-week data for the 6-week data as well, bringing the total number of participants analyzed in the nonmassage group to 18. Baseline data for participants who withdrew were not different from those who remained in the study. Table 1 shows the characteristics of the sample. We found no significant differences in sample characteristics between groups. Mothers of 69% of the infant breast-fed. Type of feeding did not correlate with any outcome variables. Mothers in the nonmassage group were more likely to discontinue antireflux medication (38% vs. 11%), but the difference was not significant.
Table 1. Baseline Characteristics of Participants.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Massage (n = 18)</th>
<th>Nonmassage (n = 18)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age (Mean (SD))</td>
<td>32.0 (4.0)</td>
<td>31.0 (4.0)</td>
<td>.32</td>
</tr>
<tr>
<td>Paternal age (Mean (SD))</td>
<td>32.0 (8.0)</td>
<td>3.0 (5.0)</td>
<td>.59</td>
</tr>
<tr>
<td>Infant age (weeks) at first treatment</td>
<td>7.3 (1.6)</td>
<td>7.6 (2.3)</td>
<td>.74</td>
</tr>
<tr>
<td>Infant weight (kg) (Mean (SD))</td>
<td>4.8 (0.5)</td>
<td>4.9 (0.6)</td>
<td>.33</td>
</tr>
<tr>
<td>Infant BMI (kg/m²) (Mean (SD))</td>
<td>14.9 (1.1)</td>
<td>14.5 (4.5)</td>
<td>.30</td>
</tr>
<tr>
<td>I-GERQ-R score at baseline (Mean (SD))</td>
<td>22.0 (4.0)</td>
<td>24.0 (4.0)</td>
<td>.44</td>
</tr>
<tr>
<td>Average daily sleep at baseline (hr)</td>
<td>9.1 (2.3)</td>
<td>8.5 (2.2)</td>
<td>.45</td>
</tr>
<tr>
<td>Maternal education</td>
<td></td>
<td></td>
<td>.90</td>
</tr>
<tr>
<td>High school</td>
<td>2 (11)</td>
<td>2 (11)</td>
<td></td>
</tr>
<tr>
<td>Some college</td>
<td>2 (11)</td>
<td>2 (11)</td>
<td></td>
</tr>
<tr>
<td>College degree (4-year)</td>
<td>6 (33)</td>
<td>8 (45)</td>
<td></td>
</tr>
<tr>
<td>Graduate degree</td>
<td>8 (45)</td>
<td>6 (33)</td>
<td></td>
</tr>
<tr>
<td>Paternal education</td>
<td></td>
<td></td>
<td>.84</td>
</tr>
<tr>
<td>High school</td>
<td>3 (17)</td>
<td>3 (17)</td>
<td></td>
</tr>
<tr>
<td>Some college</td>
<td>3 (17)</td>
<td>4 (22)</td>
<td></td>
</tr>
<tr>
<td>College degree (4-year)</td>
<td>7 (38)</td>
<td>5 (28)</td>
<td></td>
</tr>
<tr>
<td>Graduate degree</td>
<td>4 (22)</td>
<td>6 (33)</td>
<td></td>
</tr>
<tr>
<td>Father involved</td>
<td>17 (94)</td>
<td>18 (100)</td>
<td>1.00</td>
</tr>
<tr>
<td>Infant gender: male</td>
<td>12 (67)</td>
<td>11 (61)</td>
<td>1.00</td>
</tr>
<tr>
<td>Infant ethnicity</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American or Black</td>
<td>1 (6)</td>
<td>1 (6)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>2 (11)</td>
<td>2 (11)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>2 (11)</td>
<td>5 (27)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>2 (11)</td>
<td>1 (6)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>11 (61)</td>
<td>9 (50)</td>
<td></td>
</tr>
<tr>
<td>Number of siblings</td>
<td></td>
<td></td>
<td>.79</td>
</tr>
<tr>
<td>0</td>
<td>6 (33)</td>
<td>8 (44)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10 (56)</td>
<td>9 (50)</td>
<td></td>
</tr>
<tr>
<td>2 or 3</td>
<td>2 (11)</td>
<td>1 (6)</td>
<td></td>
</tr>
</tbody>
</table>

Note. BMI = body mass index; I-GERQ-R = Infant Gastroesophageal Reflux Questionnaire, Revised. SD = standard deviation.

Intervention Effects on Symptoms of GERD

We found no difference between groups in the primary outcome, GERD symptoms assessed by the I-GERQ-R, using ANCOVA adjusting for baseline value. Table 2 shows least square means of primary outcome variables and statistical outcomes. Mean scores in the massage group were 22.0 (SD = 4), 15.0 (SD = 4), and 14.4 (SD = 5) at baseline and 4 and 6 weeks, respectively; in the nonmassage group, they were 23.5 (SD = 4), 15.1 (SD = 5), and 13.7 (SD = 6), respectively.

Because reduction in irritability was thought to be a likely effect of massage, an item on the I-GERQ-R that quantified daily amount of crying was examined separately. Choices were “less than 10 minutes,” “10 minutes to 1 hour,” “more than 1 hour but less than 3 hours,” and “more than 3 hours.” At baseline, 47% of the mothers recorded between 1 and 3 hr of daily crying, while 28% reported more than 3 hr. Figure 2 displays the percentage distribution of answers to this question for each group at baseline and after 6 weeks of therapy. Statistically significant improvement was observed in the proportion of subjects crying less than 10 min (6% vs. 33%, p = .025) and in those crying less than 1 hr (17% vs. 61%, p = .0047) in the MT group, while no significant change was seen in the NMT group for crying less than 10 min (6% vs.6%, p = 1.0) or less than 1 hr (35% vs. 59%, p = .10).

The daily diary also addressed irritability, but only 14 mothers in the massage group and 13 in the nonmassage group completed the diary. Analysis compared the day before to the day after massage. Groups did not differ significantly on the average amount of infant fussing, maternal holding for pleasure, or maternal holding to prevent the infant from crying. Compared to the nonmassage group, however, infants in the massage group cried an average of 16 min per day less, and their mothers reported holding to prevent crying an average of 16 min less the day after massage. Averaged over the 6 weeks for both groups, mothers held infants 1/2 hr daily for pleasure and 2 hr daily to prevent crying.

Weight gain increased similarly in both groups. At baseline and 4 and 6 weeks, mean weight for the sample was 4.8 kg (SD = 0.6), 5.7 kg (SD = 0.7), and 6.0 kg (SD = 0.7), respectively. We also found no difference in total sleep hours between groups. Actigraphy data revealed that infants slept a mean of 8.8 (SD = 2) hr per day at baseline and 8.8 (SD = 3) hr and 9.2 hr (SD = 2) per day at weeks 4 and 6, respectively.

Intervention Effects on Salivary Cortisol

Salivary cortisol levels decreased in infants in the massage group compared to those in the nonmassage group, as shown in Figure 3. After 6 weeks of treatment, cortisol levels assessed prior to the therapy session after adjusting for baseline value was 60% lower (p = .003) in geometric mean in the massage group as compared to the nonmassage group (Table 2). In the supporting repeated measures ANOVA including data points at baseline and Weeks 4 and 6, we found a significant time-by-group interaction effect, F(6, 74.8) = 21.41, p < .0001. Specifically, post hoc analysis showed significant between-group differences in the change from baseline to Week 6 (p = .007) and from Week 4 to Week 6 (p = .016), while there was no statistically significant difference between groups in the change from baseline to Week 4. Overall infant cortisol concentrations as assessed by AUC calculated for the 10-hr samples collected by mothers at baseline and at 6 weeks decreased in the massage group while increasing in the nonmassage group. The massage group had a median decrease of 13 μg.hr/dl while the nonmassage group showed an increase of 8 μg.hr/dl. The Hodges–Lehmann point estimate of the between-group difference was 18 μg.hr/dl (95% CI −44 to 9 μg.hr/dl, with p = .11), suggesting that the massage group had a greater decrease in cortisol than the nonmassage group after therapy.

Feasibility and Fidelity

Attrition was 14.6%. Figure 1 shows reasons participants gave for withdrawal. However, mothers’ most frequent complaint about the study was the burden of completing the daily diaries, and of the first 18 participants to enroll, 7 (39%) withdrew in...
the first week of the study when completing the daily diary was required. After we decided to make completion of the daily diary optional, only 1 of the subsequent 25 (4%) participants withdrew. At the end of the study, 22% of the diaries were incomplete.

A total of 108 three-day actiwatch recordings were possible. Of these, 8 (7%) had missing data ranging from no recording (n = 2) to 2.5 days of recording. Some of the missing data were due to problems with the band on the actiwatch. In one case, a watch fell off the infant, so we made smaller bands for some of

Table 2. Analysis of Covariance for Efficacy of Intervention, Adjusting for Baseline Value.

<table>
<thead>
<tr>
<th>Endpoints assessed at Week 6</th>
<th>Massage (n = 18)</th>
<th>Nonmassage (n = 18)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-GERQ-R score</td>
<td>14.72 ± 1.27</td>
<td>13.72 ± 1.27</td>
<td>.58</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>6.02 ± 0.10</td>
<td>6.03 ± 0.10</td>
<td>.90</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>15.19 ± 0.35</td>
<td>15.80 ± 0.35</td>
<td>.23</td>
</tr>
<tr>
<td>Average daily sleep (hr)</td>
<td>9.12 ± 0.49</td>
<td>9.38 ± 0.47</td>
<td>.70</td>
</tr>
<tr>
<td>Log-transformed cortisol level*</td>
<td>0.46 ± 0.20</td>
<td>1.37 ± 0.20</td>
<td>.003</td>
</tr>
<tr>
<td>Log cortisol AUC</td>
<td>3.99 ± 0.10</td>
<td>3.98 ± 0.10</td>
<td>.93</td>
</tr>
</tbody>
</table>

Note. AUC = area under the curve; BMI = body mass index; I-GERQ-R = Infant Gastroesophageal Reflux Questionnaire, Revised.

* Cortisol level was assessed from saliva samples taken prior to the final treatment session of Week 6.

Figure 2. Amount of infant crying or fussing in 24 hr at baseline and at Week 6 in the massage and nonmassage groups. Data collected from an item on the Infant Gastroesophageal Reflux Questionnaire, Revised.

Figure 3. Salivary cortisol levels of massage and nonmassage therapy groups. Therapists collected saliva samples before treatments at baseline and at the second treatment of Weeks 4 and 6. Mothers also collected samples from their infants 3 times daily before feedings on 3 consecutive days before the first treatment at baseline and after the last treatment at Week 6. AUC = area under the curve analysis.
the infants from slightly elastic material, using velcro to attach the ends of the band together. In another case, one of these bands broke during data collection. The remainder of the missing data was due to removal because the mother thought the infant was uncomfortable.

We had no issues with infant saliva; saliva was not viscous or stringy, and we were able to obtain adequate amounts for assay. However, 10 (28%) mothers did not collect saliva from their infants.

Maternal blinding was maintained. We informed mothers of their treatment group only after all data collection was completed. At 4 and 6 weeks, we asked them whether they knew the infants’ respective treatment groups and what they liked or disliked about the study. All mothers denied having been told their infant’s treatment group, but 33% were able to guess correctly. We found no correlation between correct group identification and any of the outcomes. Mothers stated that they liked the treatment team, the time to themselves during treatments, and the feeling that their infants were being helped. They disliked completing the daily diaries, being out of the room during treatment, and not knowing what treatment the infant was receiving. Although no mothers massaged their infants, one mother in each group took her infant to a chiropractor.

Fidelity to the protocol by the therapist assessed by coders was 76% for massage therapists and 83% for nonmassage therapists. Coder agreement on therapist fidelity was 82% for massage therapists and 88% for nonmassage therapists. Deviations from the planned massage protocol occurred because of infant irritability (a) during abdominal massage or (b) when infants were turned prone for back massage. In situation (a), infants appeared to be very sensitive in the abdominal area. Therapists could do the abdominal strokes outlined in Field et al. (1996), but to avoid infant fussing, they performed these strokes very gently. In situation (b), rather than performing massage while infants were lying flat on a pad and blanket, therapists held the infants vertically on their shoulders and administered back massage with one hand. Approximately half of the complete sessions for all infants were given with the therapist holding infants vertically to minimize crying and fussing. Therapists rescheduled six massage sessions (3%) due to infant fussiness.

**Discussion**

We hypothesized that, when compared to infants who received nonmassage, infants who received massage would display (a) a greater decrease in GERD symptoms (Hypothesis 1), (b) greater weight gain (Hypothesis 2), (c) a greater increase in amount of sleep (Hypothesis 3), (d) a greater decrease in cortisol levels measured before treatments (Hypothesis 4), and (e) a greater decrease in lower overall cortisol concentrations (Hypothesis 5). Our findings did not support the primary hypothesis. GERD symptoms decreased in both groups. These results are similar to those of previous studies, which showed similar reduction of GERD symptoms regardless of whether infants were treated with antireflux medication, feeding modification, hypoallergenic formula, maternal counseling, or placebo (Jordan et al., 2006; Moore et al., 2003; Orenstein et al., 2009; Winter et al., 2010). Other studies have suggested that a variety of treatments, including extra attention given to mothers for study purposes (placebo), may lessen GERD symptoms, and a combination of treatments may be more effective than any single treatment (Neu, Corwin, Lareau, & Marcheggiani-Howard, 2011).

The therapists did not follow the planned massage protocol in approximately half of the sessions because of irritability of the infants when positioned in supine (head slightly elevated) or prone positions. Therapists successfully performed the adjusted protocol that included massage of all body areas covered in the original protocol, in 97% of sessions without the need to reschedule for fussiness. The therapists, however, were learning about the particular needs of infants with GERD during massage, and some trial and error was used in the process of adapting the protocol, which may have confounded the effect of the therapy. Interestingly, books for parents on infant massage (Heath & Banbridge, 2004; McClure, 2000) often advocate frequent abdominal massage for infants who cry excessively. Our study reveals that it is important for clinicians to impart to parents who are interested in massaging their infants with GERD that frequent abdominal massage may be ineffective and may actually increase fussiness.

Findings of other studies have indicated a reduction in irritability as mothers massaged their infants daily for 1–5 weeks (Arikan et al., 2008; Çetinkaya & Başbakkal, 2012; Scafidi & Field, 1996). In the present study, analysis of the crying item in the I-GERQ-R suggested that subjects in the massage group were less irritable after 6 weeks of therapy, although this did not affect the total GERD symptom score. The small decrease in maternal holding to prevent crying on the day following massage also suggests a decrease in infant crying the day following massage.

We found no differences in weight gain between groups. Studies in which weight increased after massage were conducted with term infants whose mothers had depression or were HIV positive (Field et al., 1996; Scafidi & Field, 1996). Neither was true of mothers in this study, who were physically healthy and whose depression and anxiety scores were well within normal ranges. Mean infant weight was in the normal range throughout the study in both groups (Hockenberry & Wilson, 2008).

Neither treatment affected sleep. Other studies in which massage enhanced sleep were done on infants from the point of birth, with the mother providing daily massage (Ferber et al., 2002). In the present study, infants in both groups, who were 2–4 months old during the treatment period, slept an average of 8–9 hr per day during the 6-week period, which is less than the average of 13-hr per day that previous studies have found in infants 3 months of age (Galland, Taylor, Elder, & Herbison, 2012). This reduced amount of sleep is similar to the findings of other research on infants with GERD, in which sleep was disrupted (Machado et al., 2013).

Our objective measure, salivary cortisol, decreased in the massage group only. Pretherapy cortisol levels decreased significantly among massage group infants from baseline to 6 weeks versus nonmassage group infants, whose levels increased. Daily...
cortisol levels also decreased over the 6 weeks, but we found no significant differences in this measure between groups. Approximately 30% of mothers in each group failed to collect saliva. Field et al. (1996) also reported decreased cortisol levels in infants of similar age (4–12 weeks) after receiving 15 min of MT twice a week for 6 weeks. It is possible that, during the months in which they have symptoms of GERD, infants experience chronic stress because of discomfort, decreased sleep, or hunger (Kleinman et al., 2006; Machado et al., 2013). Because dysregulation of the HPA system in infancy can lead to continued dysregulation later in life (Essex et al., 2011), further research on the effect of stress reduction in infants with GERD is warranted.

Feasibility

The second aim of the present study was to evaluate the feasibility of testing massage as treatment for GERD (e.g., procedure fidelity, blinding, safety, dose response, maternal acceptance, accrual, and attrition). Administration of massage was adapted to the needs of the infant, necessitating adjustments to the protocol. Blinding was maintained, but at the expense of recruitment. A typical reason mothers gave for not participating was inability to be in the room during treatment. Daily diaries yielded important information, but some mothers found its completion overwhelming. Once the daily diary requirement was removed, attrition dropped from 39% to 4%.

Limitations

Sample size was small in this pilot study. A larger sample would have allowed us to compare the impact of various massage schedules on different groups. The daily diaries may have been too complex for some mothers. A simpler version might yield a higher response rate, or information could be collected for a limited period, such as over 3 days at one or two time points during the study. In addition, the sample in this study was highly educated, and families were intact; it is unknown how massage would affect infants in higher risk families. Massage was administered only twice a week. More frequent sessions may have had a greater effect on crying and reduction in stress before 6 weeks.

Possibilities for Future Research

It is possible that stress was reduced only in the 6th week of treatment because the infants needed to adjust to the therapist. Perhaps if the therapy were administered by the mother, stress reduction would happen sooner, especially if massage were given daily. Maternal administration of massage treatments would both allow daily treatment and likely increase recruitment. The control condition would need to be a credible alternative to massage or standard treatment, such as medication only. Mother-administered massage is inexpensive, as only one or two visits from a massage therapist are necessary to teach massage to mothers. Future research is necessary to assess the potential benefits of daily maternal massage for infants with GERD. Also, a massage protocol adapted based on our experience in the present study would allow for abdominal sensitivity.

Conclusion

Findings from this study show that therapists and mothers need to be aware that the standard massage protocol for infants with GERD may need adjustment (e.g., using gentle strokes on the abdomen and providing the entire massage or back massage while holding the infant). This was the first study to examine massage as a treatment for infants with GERD. Findings from this study showing a reduction in stress levels and marginal decrease in irritability but no effect on sleep or a cluster of GERD symptoms suggest that a multifaceted treatment approach might be most effective for infants with symptoms of GERD.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding for this research was provided by the National Institute of Nursing Research, R21NR011069.

References


Neu et al.


