The Development of an Accurate Test Weighing Technique for Preterm and High-Risk Hospitalized Infants

Barbara Haase, Jeanne Barreira, Pamela K. Murphy, Martina Mueller, and Jean Rhodes

Abstract

Background: Test weighing, or weighing the infant before and after breastfeeding to assess milk intake, in which weight gain in grams is converted equally to volume of intake in milliliters, is a controversial topic in the literature. This study was initiated to identify variables that impact test weights and to develop an accurate test weighing technique for preterm and high-risk hospitalized infants.

Methods: Test weights were performed on a sample of hospitalized high-risk infants with and without leads who were bottle- or nasogastric-fed. Volume consumed was compared to weight gain to determine whether the developed technique was accurate.

Results: In each group, with or without leads, only one measure of actual intake versus test weight result was found outside the confidence limits (95%), and only one measure was found outside the clinically acceptable difference of \( \pm 5 \) g. Correlation coefficient \( (r^2) \) values of weight gain by test weight to volume of intake were 0.998 for infants without leads and 0.997 for infants with leads.

Conclusions: The data from this study support the use of this test weighing technique as an accurate, objective assessment of the measurement of breastmilk intake after a breastfeeding session, thus allowing medical decisions regarding supplementation to be based on objective data rather than inaccurate clinical indices of the quality of infant feedings at the breast.

Introduction

Test weighing, or weighing the infant before and after breastfeeding to assess milk intake, in which weight gain in grams is converted equally to volume of intake in milliliters, is a controversial topic in the literature. Since 1990, the findings of Meier and colleagues support the use of test weights for preterm and high-risk infants both in the hospital and at home. In these studies, specific models of infant scales such as the Olympic Medical (Seattle, WA) Smart Scale and the Medela, Inc. (McHenry, IL) Babyweigh scale were accurate in measuring the small weight changes that occur with feeds. In 2008 the findings of Nyquist supported the use of test weights in the management of very preterm infants as they moved towards breastfeeding competence. In contrast, two other investigators have found test weights to be inaccurate: Results from Whitfield et al. in 1981 using the Salter Trent (nondigital) scale (Trent Scales, Ilkeston, UK) and Savenije and Brand in 2006 using the Avery Berkel (Fairmont, MN) Pesa (digital) scale called into question the accuracy of test weighing. Investigators in both of these studies concluded that the inaccurate results were probably due to the imprecision of the scales.

Initiating and sustaining breastfeeding with an infant at risk for suboptimal intake are common challenges in neonatal units. Additionally, subjective assessments of breastmilk intake using timing or sucking quality during breastfeeding have been shown to be invalid methods. Objectifying milk intake can provide helpful information to guide the mother and clinician as to the appropriateness of supplementation, thus supporting breastfeeding while ensuring adequate infant intake.

The discrepancies in the literature regarding the accuracy of test weighing may be due to several factors, including the scale, the infant, or the technique. In all studies reviewed, no specific test weighing technique was described for infants requiring continuous monitoring and/or attachment to oxygen tubing during the procedure. Therefore, the purpose of our study was to identify variables that impact test weights and to develop an accurate test weighing technique for preterm and high-risk hospitalized infants.
Materials and Methods

Development of test weighing technique

Pre- and post-test weights were obtained in the intensive care nurseries of a tertiary care medical center to develop an accurate test weighing technique to assess breastmilk intake while breastfeeding. This initial phase included a convenience sample of 19 preterm infants, who were either bottle- or nasogastrically fed breastmilk, fortified breastmilk, or formula. The investigators measured the volume of intake during the feeding session. As infants were weighed, the investigators carefully observed for variations in technique or circumstances that influenced the accuracy of the weights. These variables included: the scale; infant movement; positioning of the infant on the scale; monitor lead wires/oxygen tubing/clothing touching the edge of the scale; tension on the lead wires/tubing; and inconsistent positioning of the scale.

Insensitive water loss, could also impact test weight accuracy. Previous studies\textsuperscript{10,11} described dressing and wrapping infants prior to weighing as a process of standardizing clothing and preventing insensitive water loss by the cutaneous route. Woolridge et al.\textsuperscript{7} suggested insensitive water loss can be insignificant over the course of individual feedings. And, while the fat content of milk can minimally alter the conversion of weight in grams to volume in milliliters by up to 3%, no correction for the density of milk is necessary.\textsuperscript{7} Meier et al.\textsuperscript{11} agreed and further explained that the increased weight of higher-density, higher-fat milk may cancel out the small decrease in weight from insensitive water loss.

Three digital scales were initially used for this study: the Olympic Smart Scale, the Scaletronic (Carol Stream, IL) pediatric scale model 4800, and the Medela Babyweigh scale. All three scales were calibrated for accuracy within 2 g; however, each type of scale presented different scenarios in the process of weighing an infant. The Olympic Smart Scale has a high-walled bucket style platform, making it difficult for investigators to keep any connected lead wires and/or tubing off the edge of the scale. The Scaletronic pediatric scale has a flat platform with open ends, allowing for better management of the lead wires and tubing. However, measurements varied 10 g or more depending on the placement of the infant on the scale. The Medela Babyweigh scale has a short, slightly upward curving edge that does not interfere with management of the lead wires/tubing. While the Medela Babyweigh scale was used for the blinded portion of this study because it was conducive to controlling certain variables affecting test weights, each institution should evaluate the accuracy of its scales in the clinical setting to determine the suitability with test weighing.

In the process of developing a test weight technique we found that infant movement such as flailing arms and legs can cause inaccurate results; therefore, we recommend tightly swaddling the infant prior to weighing. We found clothing or blankets draping over the scale can also cause weight variation. Any burp cloths, diapers, or other articles related to the infant should be included in both the pre- and post-weights.

The last factor contributing to weight variation is the attachment of an infant to monitors or oxygen. This factor has been previously mentioned by Meier and colleagues\textsuperscript{10,11,13}; however, no method for controlling weight variation has been described. Tension or dangling of tubing, wires, or the lead connection point or holding them above the infant—all common mistakes noted in the clinical area—can lead to significant weight variation. It is important to put the cart holding the scale as near as possible to the infant’s crib and in the same place for the pre- and post-weights. Taping the lead connection point and/or tubing to the crib prevents lead wires/tubing from touching the edge of the scale while avoiding unwanted tension. When a post-feeding weight was determined for the infants with leads, the scale was turned off while the researcher repositioned the infant on the scale and retaped the leads. If the scale was not turned off the post-feeding weight would be calculated by the scale while the infant and leads were being resituated. This would have lead to an incorrect post-feed weight. Furthermore, prior to the study, nurses told the researchers they could not manage the process of test weighing without assistance in infants with leads. The procedure described in this study allows a single nurse to manage all aspects of the procedure.

A technique combining the above elements for test weighing was developed and then tested in blinded studies on infants with leads and without leads. In this research article, the term “infants without leads” refers to infants not requiring continuous monitoring or infants who could be disconnected from monitors for short periods of time (for the weighing) and were not attached to oxygen tubing. The term “infants with leads” refers to infants requiring continuous monitoring who could not be disconnected from monitors and/or were attached to oxygen tubing.

Inter-rater reliability for the test weighing technique was established among four investigators by having each investigator perform several test-weights with results falling within a variance of 5 g. In addition, investigators repeated each measure twice to verify accuracy between measures of at least 5 g.

Test weighing technique: Infants without leads

A convenience sample of 10 neonatal intensive care (NICU) infants between 31 and 37 weeks adjusted gestational age, weighing between 1,594 and 3,220 g, were either bottle- or nasogastrically fed breastmilk, fortified breastmilk, or formula with volumes between 8 and 77 mL measured by volu-feeder or syringes. All nurses feeding infants were experienced NICU nurses. The researchers were blinded to the volume of milk consumed during the feeding until the post-weight was calculated. Infants were diapered and clothed, and the monitor alarm was suspended. The monitor lead wires attached to the infant were disconnected from the distribution block of the trunk cable (the point at which the monitor lead wires were connected to the cable leading to the monitor), referred to from this point on as the “lead connection point.” Lead wires were placed on the infant’s abdomen, and the infant was swaddled tightly in a blanket.

The Medela Babyweigh scale was placed on a level cart, and the scale was turned on in the grams mode and zeroed. The infant was placed on the scale in a supine position, ensuring that all blankets, clothing, burp cloths, the infant’s extremities, and/or monitor leads were not touching the side of the scale. The pre-weight was obtained twice, and the average weight was recorded. If a discrepancy of more than 5 g between the weights existed, the weight was repeated un-
To determine infant intake based on test weights, the average pre-weight was subtracted from the average post-weight. The volume consumed during the feeding was obtained from the nurse. The difference between test weight data and the volume consumed was compared.

Statistical methods

In a first step, standard scatter plots of weight gains (difference between averaged pre-weight and averaged post-weight) versus volume intake (actual oral or nasogastric intake) were developed as a first estimate of the degree of agreement between the two measurements. Further, Pearson correlation coefficients were determined for weight gains for both the group with and the group without leads. Means and ranges of actual intake were calculated for both groups.

Then Bland-Altman plots were used to assess the level of agreement between weight gain measured using the standardized weighing procedure and feed volume taken by the infant. Because of variation inherent in any measurement method, no two measurements will result in exactly equal values, and neither of two methods will produce the absolute true value. However, if the two methods are considered to be “in agreement” (i.e., differ by not more than a prespecified amount), the new method can be used in place of the established one. Values of 0 for the difference between the two methods to determine weight gain would mean that the methods are identical. By showing the deviation of individual measures of the two methods from zero and therefore the spread from the ideal (a difference of zero) across the entire data set, the Bland-Altman plot provides an additional measure of accuracy. Also, the mean of the two measures can be used as the best estimate to compare a new method with an established one. The degree of clinically acceptable difference between the two methods used in this study was determined to be ±5 g. All statistical tests were performed using SAS statistical software version 9 (SAS Institute Inc., Cary, NC).

Results

Actual intake

Mean actual intake among infants without leads was 43.4 ± 17.9 g, ranging from 20 to 80 g. For the infants with leads, the actual volume intake ranged from 8 to 77 g, with a mean of 42.3 ± 20.0 g.

Agreement between pre-weights and post-weight pairs

Of the 20 pairs of weights obtained on 10 infants without leads (one pair = either the two pre-weights or the two post-weights), 17 pairs were identical, two pairs differed by 2 g, and one pair differed by 4 g.

Of the 24 pairs of weights obtained on 12 infants with leads (one pair = either the two pre-weights or the two post-weights), 20 pairs were identical, three pairs differed by 2 g, and one pair differed by 4 g.

Tests of association of test weights and actual intake

Weight gain (in g) versus volume intake (in mL), for infants weighed without leads (n = 10) (Fig. 1A) and infants weighed with leads (n = 12) (Fig. 1B), was plotted on graphs
showing the linear correlation of the two measurements. The plotted line represents the line of equality, on which all points would fall if the two measurement methods gave the exact same result every time. For both plots all points lay within a relatively narrow distance from the line of equality, which was confirmed by the Pearson correlation coefficient ($r^2$) of 0.998 for infants without leads and $r^2$ of 0.997 for infants with leads.

**Agreement between test weights and actual intake**

The levels of agreement between weight gain measured using a standardized weighing procedure and known volume intake for infants weighed without and with leads are shown in Figure 2. Differences between weight gain (in g) and volume intake (in mL) are plotted on the vertical axes, while means of weight gain and volume intake are plotted on the horizontal axes. The shaded area reflects the predetermined clinically acceptable difference ($\pm 5$ g) between weight gain measured by the standardized weighing procedure and known volume intake.

For infants weighed without leads, 95% limits of agreement (i.e., mean difference $\pm [1.96 \times s]$) were $-6.03$ and $3.83$. The 95% confidence interval for the bias (mean difference) was $-2.90$ to $0.70$, the 95% confidence interval for the lower limit of agreement was $-9.14$ to $-2.92$, while the 95% confidence interval for the upper level of agreement was $0.72$ to $6.94$.

For infants weighed with leads, 95% limits of agreement (i.e., mean difference $\pm [1.96 \times \text{mean}$]) for the two measurements were $-3.61$ and $5.77$. The 95% confidence interval for the bias (mean difference between weight gain and volume intake) was $-0.44$ to $2.60$, the 95% confidence interval for the lower limit of agreement was $-6.24$ to $-0.97$, while the 95% confidence interval for the upper level of agreement was $3.14$ to $8.40$.

In each group only one point falls outside the confidence limits, and only one point falls outside the clinically acceptable difference of $\pm 5$ g.

**Discussion**

The purpose of this study was to develop an accurate test weighing technique for use with preterm and high-risk hospitalized infants. With 95% confidence, the test weighing technique developed by the investigators in this study demonstrated accuracy within 5 g of the actual volume intake by the infant fed breastmilk, fortified breastmilk, or formula via a nasogastric tube or orally with a bottle. These results support the conclusions drawn from previous studies that test weights are an accurate method for measuring breastmilk intake during a breastfeeding session when the procedure is performed correctly and consistently.\(^{10,11,13}\)

While both Whitfield et al.\(^{8}\) and Savenije and Brand\(^{9}\) described test weighing as an imprecise method of assessing breastmilk intake from a breastfeeding session, neither study described any consistent method for assessing test weights or controlling for variables that can contribute to inaccurate test weighing results. Meier et al.\(^{13}\) described the importance of using consistent, controlled methods to establish whether a measurement technique is accurate.

In establishing the described test weighing technique, several confounding variables were controlled, including minimizing infant movement, eliminating clothing or blankets from draping over the side of the scale, eliminating tension on the leads, and maintaining consistent positioning of the leads and/or tubing. In addition, infants’ clothing, diaper, hat, and burp cloths were included in both the pre- and post-test weight. Reliability of the test weight procedure was established among all investigators who performed the test weighing technique, which was not done.
in the studies done by Whitfield et al.\textsuperscript{8} and Savenije and Brand.\textsuperscript{9}

\textit{Limitations}

This research is limited by a small sample size: 19 infants in the initial phase and 22 in the blinded phase. The investigators did not test the technique on infants weighing less than 1,594 or more than 3,220 g or on infants with intravenous fluids. Neither infant gender nor race was considered.

In addition, few measures of infant intake were less than 20 mL. Since many preterm infants consume low volumes at breast, further study should evaluate this technique in in-
Fants with intake volumes less than 20 mL. The length of time for each feeding was not recorded, although pre- and post-weights were measured immediately before and after each feed. Future studies could address these issues as well as test the technique used in this study with other scales. This study has articulated and evaluated one accurate test weighing technique, but with further studies, other techniques may also be developed.

Conclusions

The data from this study support the use of this test weighing technique as an accurate, objective assessment of the measurement of breastmilk intake after a breastfeeding session. This technique can be implemented into clinical practice to encourage breastfeeding and provide accurate knowledge of a preterm or high-risk hospitalized infant’s oral intake at the breast so that over- or undersupplementation can be avoided. Additional research should include a larger sample size, demographic factors such as gender or race, measurements of volume intake under 20 mL, and the use of other scales to either replicate or develop new accurate test weighing techniques for preterm and high-risk hospitalized infants.

A video of the test weight procedure can be found on the website <http://www.muschealth.com/women/lactation/pumping.htm>.

Acknowledgments

We would like to thank Lynne Nemeth, Ph.D., Suzanne O. Gresle, Susan Newman, Ph.D., and Jennie Ariail, Ph.D., for their editorial support.

Disclosure Statement

No competing financial interests exist.

References


Address correspondence to:
Jean Rhodes, Ph.D.
Lactation Consultation Service
Medical University of South Carolina
169 Ashley Avenue
P.O. Box 250347
Charleston, SC 29425

E-mail: rhodesje@musc.edu