Use of the Broselow Tape May Result in the Underresuscitation of Children

Carolyn T. Nieman, MSN, ACNP, Christopher F. Manacci, MSN, ACNP, Dennis M. Super, MD, MPH, Charlene Mancuso, RN, MPA, William F. Fallon, Jr., MD, MBA

Abstract

Objectives: The purpose of this study was to determine the concordance of the Broselow tape with the measured heights and weights of a community-based population of children, especially in light of the increase in obesity in today's children.

Methods: The authors examined more than 7,500 children in a cross-sectional, descriptive study in two different cohorts of children to compare their actual weight with their predicted weight by a color-coded tape measure.

Results: In all patients, the percent agreement and κ values of the Broselow color predicted by height versus the actual color by weight for the 2002A tape were 66.2% and 0.61, respectively. The concordance was best in infants, followed by school-age children, toddlers, and preschoolers (κ = 0.66, 0.44, 0.39, and 0.39, respectively; percent agreement, 81.3%, 58.2%, 60.7%, and 64.0%, respectively). The tapes accurately predicted (within 10%) medication dosages for resuscitation in 55.3%–60.0% of the children. The number of children who were underdosed (by ≥10%) exceeded those who were overdosed (by ≥10%) by 2.5 to 4.4 times ($p < 0.05$). The tapes accurately predicted uncuffed endotracheal tube sizes when compared with age-based guidelines in 71% of the children, with undersizing (≥0.5 mm) exceeding oversizing by threefold to fourfold ($p < 0.05$).

Conclusions: The Broselow tape color-coded system inaccurately predicted actual weight in one third of children. Caregivers need to take into consideration the accuracy of this device when estimating children's weight during the resuscitation of a child.

ACADEMIC EMERGENCY MEDICINE 2006; 13:1011–1019 © 2006 by the Society for Academic Emergency Medicine

Keywords: resuscitation, pediatric, Broselow, color-coded tape, length-based tape

The practice of caring for critically ill and injured children requires meticulous attention to detail and a high level of diligence when determining dosages of medications and selecting appropriately sized equipment. Most of these important decisions are based on an accurate estimate of a child's weight. While non–critically ill children can be weighed on a scale, this method is impractical for unstable patients because of the urgency of their condition.

Practitioners have come to rely on the use of measurement tools where the child's length is used to estimate weight. One tool used to estimate the weight of a pediatric patient is the Broselow color-coded system.[1] This measuring device divides children into weight categories based on length. Each category is assigned a color that corresponds to lists of appropriately sized equipment for that patient. Likewise, the color-coded system also provides drug dosages based on the estimated weight of the child.[2]

Over the past three decades, the prevalence of obesity in American children has increased at an alarming rate. The National Health and Nutritional Examination Survey (NHANES) III cohort study suggested that approximately 14% of U.S. children are obese, as defined by body mass index.[3,4] Faced with this alarming rate of obesity,

From the National Flight Nurse Academy, Bolton School of Nursing (CTN, CFM), Division of Pediatric Critical Care, Department of Pediatrics, MetroHealth Medical Center (DMS), and Division of Trauma, Critical Care, Burns and Life Flight (CM), MetroHealth Medical Center, Case Western Reserve University, Cleveland, OH; and Akron City Hospital (WFF), Akron, OH.

Received December 29, 2005; revisions received June 7, 2006, and June 11, 2006; accepted June 15, 2006.

Presented in part at the Aeromedical Transport Conference meeting, Salt Lake City, UT, November 2002.

Supported in part by the General Clinical Research Center at Case Western Reserve University (National Institutes of Health grant no. MO1-RR00080).

Address for correspondence and reprints: Carolyn T. Nieman, MSN, ACNP, School of Nursing, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106. Fax: 216-368-3542; e-mail: carolyn.nieman@case.edu.

doi: 10.1197/j.aem.2006.06.042
ISSN 1069-6563
PII ISSN 1069-6563583

a length-to-weight estimate system such as the Broselow tape should be reevaluated.

A review of domestic and international literature reveals little scientific data related to the efficacy of length-based systems of resuscitation. However, using a length-based system in determining the resuscitation needs of children has proven more successful than when no calculation aids are used.[5] In 1988, Lubitz et al. reported that the Broselow tape is a "simple and accurate method of estimating pediatric weights and drug dosages," eliminating the need for calculations. However, an error rate exceeding 10% of the actual weight was noted in 40.3% of the children measured.[2] In 2002, Luten reported that the Broselow tape was superior to clinical estimates of weight. However, the author acknowledged, "resuscitation aids provide drug doses only as accurate as the weight estimation used to access them."[6,7]

A study of methods to estimate weight in children conducted in 2002 in Victoria, Australia, concluded that length-based estimations were superior when compared with other calculation aids in the care of children.[8] Conversely, Hofer et al. reported that "tracheal tube size selection using the Broselow tape was not superior to the age-based formula."[9]

Given the current state of controversy and the alarming rate of childhood obesity, we determined a need for a critical and comprehensive investigation of the Broselow system in meeting the interventional needs of critically ill and injured children. We hypothesized that given the current rate of childhood obesity, the Broselow color-coded resuscitation system would underestimate actual weight based on length in a significant percentage of children.

METHODS

Study Design

This was a cross-sectional, descriptive study comparing actual and Broselow tape–predicted weights in two cohorts of children less than 12 years of age. The Institutional Review Board for MetroHealth Medical Center approved this study. This study was also approved by the administration of each participating school in accordance with their policies regarding human investigations. Written informed consent was obtained from the parents, coupled with written assent from children older than 7 years.

Study Setting and Population

Two different cohorts of children were assembled from northeast Ohio to compare their actual weight with their predicted weight by a color-coded tape measure (Broselow Pediatric Emergency Tape; Armstrong Medical Industries, Lincolnshire, IL). One cohort was recruited from five suburban elementary schools (suburban cohort), and the second cohort was recruited from an urban, general pediatric clinic (urban clinic cohort). Patients were excluded if their height fell outside of the tape measure's range.

Study Protocol

Suburban Cohort. In May 2001, all children from five elementary schools in the eastern and western suburbs of Cleveland, Ohio, were enrolled consecutively. At the time of enrollment, an investigator measured the children's length and weight. All measurements were performed with shoes removed and with children wearing light apparel appropriate for late spring weather. Standing length was recorded both in centimeters and in the corresponding Broselow color (1998 version). A Detecto medical scale (Detecto Scales, Webb City, MO) measured the children's weight in kilograms. The auditors of Geauga and Cuyahoga Counties certified the accuracy of the scales, which were within 0.25 kg. A second investigator independently remeasured every tenth child. To reduce transcription errors, a third investigator entered the data into a computerized spreadsheet at the time of data collection.

Urban Clinic Cohort. Since 1999, the pediatric clinic at MetroHealth Medical Center has utilized a computerized medical record for all outpatient visits (Epi Care; Epic Systems Corp., Madison, WI). MetroHealth Medical Center (Cleveland, OH) is a county-owned facility that provides both tertiary care (750 beds, Level 1 adult and pediatric trauma center) and primary care to the residents of Cuyahoga County. The general pediatric clinic primarily serves the inner city. Using the computerized medical record system, the information service department identified the first well-child visit of 2000 for all children seen at this general pediatric clinic who were between 2 weeks and 11 years of age. From these medical records, the information service department abstracted the patient's height, weight, race, gender, age, and date of visit.

In this pediatric clinic, children younger than 1 year were weighed wearing only a diaper using a Scale Tronix 4800 (accuracy, $\pm$0.003 kg; New York, NY). Children 1 year of age or older were weighed in their stocking feet wearing ambient clothing using a Scale Tronix scale (accuracy, $\pm$0.100 kg). The height of the children younger than 1 year or those unable to stand was measured in the supine position using an Infantometer measuring device (accuracy, 3 mm; Perspective Enterprise, Portage, MI). For children able to stand, a Stadiometer was used to measure the standing height (accuracy, 0.375 cm; Harpenden Stadiometer; Holtain, Crymych, Great Britain).

To assess the urban clinic cohort data for possible data entry errors, the investigators reviewed a subset of medical records to determine if the child's anthropometric measurements for that clinic visit were consistent with their respective growth curve. The weight color categories were assigned a number ranging from 1 to 9 (gray, 1; green, 9). Based on the difference between the actual weight color category minus the weight color category predicted by height, the following sampling strategy was used: >2 categories (weight greater than height; all charts reviewed), 2 categories (every tenth chart), 1 category (every 40th chart), -1 category (every tenth chart), and ≤ -2 categories (weight less than height; all charts reviewed).

Emergency Tape Measure Systems

This study evaluated both the 1998 and the 2002A versions of the Broselow Pediatric Emergency Tape (Table 1). We chose to study both tapes because we

Table 1
Broselow Height-based System for Predicting Weight

Color Zone	1998 Tape	2002A Tape
Gray/infant		
Height (cm)	46.1–60.69	42.20–60.79
Predicted weight (kg)	3–5	3–5
Pink		
Height (cm)	60.70–67.69	60.80–67.79
Predicted weight (kg)	6–7	6–7
Red		
Height (cm)	67.70–75.29	67.80–75.29
Predicted weight (kg)	8–9	8–9
Purple		
Height (cm)	75.30–84.99	75.30–85.09
Predicted weight (kg)	10–11	10–11
Yellow		
Height (cm)	85.00–98.19	85.10–97.79*
Predicted weight (kg)	12–14	12–14
White		
Height (cm)	98.20–110.69	97.80–110.19*
Predicted weight (kg)	15–17	15–17
Blue		
Height (cm)	110.70–122.39	110.20–121.89*
Predicted weight (kg)	19–22	19–22
Orange		
Height (cm)	122.40–137.29	121.90–133.69*
Predicted weight (kg)	24–30	24–28*
Green		
Height (cm)	137.30–143.6	133.70–146.59*
Predicted weight (kg)	32–34	30–36*

* Differences between 1998 and 2002A tapes.

have found in our experience that many health care systems still use the older version. In addition, the evaluation of both tapes provides the clinician with information to determine if the improvement in concordance warrants the replacement of the older tape.

The 2002A tape is based on NHANES III data, whereas the 1998 version is derived from NHANES II data. The major differences in the 2002A tape versus the 1998 tape are the reduction of the yellow zone (12–14 kg) by approximately 0.5 cm, the decrease in the orange zone (24–28 kg) by placing the 30-kg subgroup into the green zone, and an increase in the green zone (30–36 kg) from 34 to 36 kg while the overall length of the tape is reduced by almost 8.5 cm (Table 1). In addition, the medication dosages for the 2002A version are reported as one set of dosages per color based on the midrange weight for that color zone. The older version reports doses in increments of 1 kg (3–20 kg) or 2 kg (22–34 kg) for the predicted weight.

The length corresponding to each color zone was measured by taping the Broselow tape flat on all four edges to a hard surface. Each tape was independently measured on three separate days with a 61-cm stainless steel ruler graduated in 1-mm increments.

The patient's length was then recoded into the respective color and drug dosage zones. In the 1998 version, the drug dosage zones end at 34 kg, which is 11 cm less than the end of the green zone. Hence, it was decided to analyze the 1998 tape ending at the 34-kg zone.

For drug dosages, both tapes via the child's height will produce an estimated weight, which can be directly compared with the patient's actual weight. To compare color zone by height versus color zone by actual weight, the actual weight would need to be recoded into a color zone. Each color zone has a corresponding weight range (2002A: blue zone, 19–22 kg; orange zone, 24–28 kg; green zone, 30–36 kg). If the child's weight falls within that range, their weights can be directly recoded into the respective color zone. Because some children's weights do not reside within the color zone ranges (22.01–23.99, 28.01–29.99), a sensitivity analysis was conducted with the following a priori definitions (Table 2). The "liberal" definition placed the entire lower segment (22.01–23.99) into the next highest color zone (liberal orange: 22.01–23.99 plus 24.00–28.00). The "conservative" definition placed the entire upper segment into the respective color zone (conservative orange: 24.00–28.00 plus 28.01–29.99). The "moderate" definition divided both segments in half and incorporated the contiguous halves in the respective color zone (moderate orange: 23.00–23.99 plus 24.00–28.00 plus 28.01–28.99). A prior evaluation of the Broselow tape system only used the moderate definition.[6] The definition of these zones can affect concordance. If the study population is more obese, the liberal definition will show greater discordance with the Broselow system. The Results and Discussion sections will focus on the moderate definition, with the liberal and conservative definitions presented in the Tables.

To determine uncuffed endotracheal size, the age-based formula ([age/4] + 4) was used for children older than 2 years. The calculated size was then rounded to the nearest half tube size (i.e., 4.3 rounded to 4.5). For children younger than 2 years, the uncuffed endotracheal tube size was abstracted from the Pediatric Advanced Life Support (PALS) guidelines (0–9 months, 3.5; 9.1–18 months, 4.0; 18.1–24 months, 4.5).[10]

Data Analysis

The following descriptive statistics were used in reporting the data: mean with the standard deviation (SD) (parametric, interval data), median with the intraquartile range (ordinal and nonparametric interval data), and percentages with the corresponding 95% confidence interval (categorical data). The following statistical tests were used to show differences between the two cohorts:

Table 2
Definition of Color Zone by Actual Weight

Color Zone (weight/tape)*	Liberal	Moderate	Conservative
Gray (3–5)	3.00–5.00	3.00–5.49	3.00–5.99
Pink (6–7)	5.01–7.00	5.50–7.49	6.00–7.99
Red (8–9)	7.01–9.00	7.50–9.49	8.00–9.99
Purple (10–11)	9.01–11.00	9.50–11.49	10.00–11.99
Yellow (12–14)	11.01–14.00	11.50–14.49	12.00–14.99
White (15–18)	14.01–18.00	14.50–18.49	15.00–18.99
Blue (19–22)	18.01–22.00	18.50–22.99	19.00–23.99
Orange 1998 (24–30)	22.01–30.00	23.00–30.99	24.00–31.99
Orange 2002A (24–28)	22.01–28.00	23.00–28.99	24.00–29.99
Green 1998 (32–34)	30.01–34.00	31.00–34.99	32.00–35.99
Green 2002A (30–36)	28.01–36.00	29.00–36.99	30.00–37.99

* Weight reported in kilograms for each color zone on the tape.

t-test (separate or pooled variance depending on the *f* ratio of the variances) for parametric, interval-level data and chi-square test for categorized data. Analysis of variance was used to determine the differences among age categories for interval level data. If the analysis of variance was statistically significant ($p < 0.05$), post-hoc t-tests between the respective groups were conducted with the Bonferroni correction for reducing test-wise error. Similarly, for categorical variables with tables greater than 2×2 (i.e., accuracy rate by age category), the data were first analyzed with the chi-square test. If statistically significant, the tables were partitioned into 2×2 tables (accuracy rate by infant/toddlers) and analyzed with post-hoc chi-square tests with the Bonferroni correction for reducing test-wise error.[11] Concordance between the actual weight zones and the weight zones predicted by height was assessed by percent agreement (PA) and the κ statistic for categorical data and by the Bland–Altman technique for interval data.[12] The κ statistic removes perchance agreement from the PA. κ interpretation for the degree of concordance is defined as follows: excellent, κ of 0.75–1.00; fair to good, κ of 0.40–0.74; poor, κ of <0.40. Acceptable concordance is defined as a κ value ≥0.40.[13] For the Bland–Altman technique, "accuracy or bias" is calculated as the mean of the difference between the two values (i.e. predicted weight minus actual weight), whereas precision is represented as the SD of that difference.[12] The data were analyzed using SPSS version 13.0 (SPSS Inc., Chicago, IL). Statistical significance was defined a priori as a two-tailed p value of <0.05.

RESULTS

Of the 8,265 eligible children, 48 were excluded because their heights were less than the threshold for both tapes (46.1 cm). An additional 546 were excluded from the 1998 tape analysis and 404 from the 2002A tape analysis because their heights exceeded the upper limit of the tape (height >143.6 cm and height >146.6 cm, respectively). Therefore, there were a total of 7,671 (urban clinic cohort, 6,732; suburban cohort, 939) and 7,813 (urban clinic cohort, 6,815; suburban cohort, 998) children included in the analysis for the 1998 and 2002A tapes, respectively (Table 3). The patient characteristics for the 7,813 patients used for the 2002A tape analysis are reported in Table 3. In the urban clinic cohort, race and gender were similar among the four age categories. However, the school-age urban clinic cohort had a higher percentage with a body mass index ≥90 percentile than the preschool group ($p < 0.05$). For the school-age children, the suburban cohort had fewer African American and obese children (body mass index ≥90 percentile) than the urban clinic cohort ($p < 0.05$).

With regard to the accuracy of the data in the suburban cohort, there was complete agreement between both sets of measurements (every tenth child remeasured by a second investigator) for all the children. In the urban clinic cohort, the review of the height and weight growth curves detected no data entry errors except when the differences (color zone by actual weight minus color zone predicted by height) were >2 zones (*n* = 1, weight off by 10 kg) and ≥−2 zones (*n* = 14, recording centimeters as inches or kilograms as pounds). The above corrections in the data set were made before the analysis. For the calibration measurement of the tapes with the steel ruler, all repeat measurements were within 1 mm.

For all patients, the PA with the Broselow color predicted by height versus weight was the same for both tapes at 66%, with κ values of 0.60 and 0.61 (Table 4, moderate definition, 1998 and 2002A tapes, respectively). When stratified by age, the PA and κ values were the highest for the infants (PA, 81.3%; κ = 0.66 for both tapes; $p < 0.05$). In toddlers and preschool children, the PA decreased to the low 60s ($p < 0.05$), with κ values in the range of 0.37–0.39. In school-age children, the PA and κ values were higher with the 2002A versus the 1998 tape (PA, 58.2% vs. 56.2%; $p < 0.05$; κ = 0.44 vs. 0.39, respectively). Even though the overall κ score of the 2002A tape was 0.61, only two of the four age groups (infants, school-age children) had a κ value ≥ 0.40 (Table 4, moderate definition).

In the two school-age cohorts, the PA and κ values for the 1998 tape were higher in the suburban cohort versus the urban clinic cohort (Table 5; PA, 60.5% vs. 54.0%; $p < 0.05$; κ = 0.40 vs. 0.38). Similar findings were noted with the 2002A tape (Table 5).

The overall accuracy of the 2002A versus the 1998 tape in predicting actual weight improved with a reduction in the percent difference from −5.6% to −3.9% ($p < 0.05$; Table 6). When stratified by color zone, the 2002A tape outperformed the 1998 tape in six of nine color zones ($p < 0.05$), with the greatest improvement in percent difference occurring in the orange zone (−7.3 to −1.5; $p < 0.05$). The 1998 tape was better in two zones (infants, purple zone; $p < 0.05$), with both tapes being similar for the

Table 3
Patient Characteristics (2002A Tape)

	Urban Clinic Cohort				Suburban Cohort
	Infant (*n* = 2,249)	Toddler (*n* = 1,403)	Preschool (*n* = 1,224)	School Age (*n* = 1,939)	School Age (*n* = 998)
Age range (mo)	0.1–11.9	12.0–35.9	36.0–59.9	60.0–144.0	60.0–144.0
Age (mo), mean ± SD	2.8 ± 3.2	21.8 ± 7.0	47.6 ± 7.2	92 ± 22.4	95.9 ± 18.8
Female (%)	51.8	48.3	49.3	48.3	48.6
African American (%)	31.2	29.7	33.1	31.5*	2.0*
Body mass index			16.3 ± 1.8	17.7 ± 3.6*	16.9 ± 2.5*
Body mass index % ≥90th percentile)			20.9†	24.0*,†	14.1*

* p < 0.05, urban school age vs. suburban school age.
† p < 0.05, urban preschool vs. urban school age.

Table 4
Concordance of Color Zone by Height versus Color Zone by Actual Weight

	1998 Tape (*n* = 7,671)		2002A Tape (*n* = 7,813)	
	κ	Percent Agreement	κ	Percent Agreement
Moderate definition				
Total	**0.60**	65.6	**0.61**	66.2
Infant	**0.66**	81.3*	**0.66**	81.3*
Toddler	0.39	60.9*,†	0.39	60.7*
Preschool	0.37	63.6*,‡	0.39	64.0*,‡
School age	0.39	56.2*,†,‡,§	**0.44**	58.2*,‡,§
Liberal definition				
Total	**0.52**	58.8§	**0.54**	60.0 §
Infant	**0.57**	75.5*	**0.57**	75.5*
Toddler	0.23	50.2*,†	0.23	50.2*,†
Preschool	0.28	57.8*,†,‡	0.30	58.4*,†,‡
School age	0.30	49.7*,‡,§	0.38	53.6*,‡,§
Conservative definition				
Total	**0.63**	67.6	**0.63**	67.3
Infant	**0.66**	82.3*	**0.66**	82.3*
Toddler	**0.42**	62.5*,†	**0.43**	62.9*,†
Preschool	**0.44**	67.8*,†,‡	**0.45**	68.3*,†,‡
School age	**0.42**	58.3*,†,‡,§	**0.44**	57.6*,†,‡,§

Boldface indicates a κ value in the fair to good range (0.40–0.74).
* $p < 0.05$, infants vs. toddlers, preschool, and school age.
† $p < 0.05$, toddler vs. preschool and school age.
‡ $p < 0.05$, preschool vs. school age.
§ $p < 0.05$, 1998 vs. 2002A tape.

Table 5
Concordance of Color Zone between Urban and Suburban Schoolchildren

	Urban Clinic Cohort		Suburban Cohort	
	κ	Percent Agreement	κ	Percent Agreement
Moderate definition				
1998 tape	0.38	54.0	**0.40**	60.5*
2002A tape	**0.42**	55.5	**0.47**	63.2*
Liberal definition				
1998 tape	0.29	47.3	0.31	54.4*
2002A tape	0.36	51.0	**0.41**	58.8*
Conservative definition				
1998 tape	**0.41**	56.6	**0.42**	58.3*
2002A tape	**0.42**	55.5	**0.45**	61.7*

Boldface indicates a κ value in the fair to good range (0.40–0.74).
* $p < 0.01$, between urban and suburban cohorts.

red zone. For children in the green zone who were ≤143.6 cm in length, the percent difference improved with the 2002A tape (−5.5 vs. −10.2; $p < 0.05$). For children with a height greater than 143.6 cm but less than 146.6 cm (2002A tape, upper green zone), the percent difference was threefold higher than those in the lower portion of the green zone (−17.7 vs. −5.5; $p < 0.05$). With regard to "precision," the SDs for percent differences with the 1998 and 2002A tapes were similar (SD, ±11.9% and ±13.1%, respectively). When stratified by color zone, the SDs were also similar for the two tapes, with SDs being slightly higher in all color zones for the 2002A tape (range in differences between SD for all color zones: 0.2% [infant zone] to 2.4% [orange zone]) (Table 6).

In recoding the percent difference in weights from Table 6 into accurate (within 10%) and inaccurate (≥10% difference), the overall accuracy rate for estimating actual weight was higher with the 1998 tape (60.0% vs. 55.3%; $p < 0.05$) (Table 7). When stratified by age categories, the accuracy rate was higher in the 1998 tape for all four age categories ($p < 0.05$). Both tapes were two to five times more likely to underestimate rather than to overestimate drug dosages for the four age categories ($p < 0.05$). The accuracy rate (within 10%) was higher in the suburban cohort in contrast to the school-age urban cohort (2002A tape, 54.8% vs. 49.3%; $p = 0.002$).

The accuracy of the Broselow tape in estimating uncuffed endotracheal tube size in comparison with the age-based PALS guidelines was 71%, with a κ of 0.65–0.66 (Table 8). The accuracy of the 2002A tape was best in infants, followed by toddlers, preschoolers, and school-age children ($p < 0.05$). In addition, the accuracy for school-age children was better for the 2002A tape ($p < 0.05$). The tape system was more likely to select a

Table 6
The Bias and Precision of the Broselow Tapes in Predicting Actual Weight

Broselow Color (No. for 2002A Tape)	Predicted Weight (kg) (2002A Tape)	Weight Difference (1998 tape)	Weight Difference (2002A Tape)	Percent Difference (1998 Tape)	Percent Difference (2002A Tape)
All (7,813)	15.6	−1.46 ± 3.60	−1.01 ± 3.68*	−5.6 ± 11.9	−3.9 ± 13.1*
Infant (1,545)	3.5	−0.16 ± 0.49	−0.23 ± 0.50*	−2.8 ± 12.9	−4.3 ± 13.1*
Pink (330)	6.5	−0.35 ± 0.72	−0.28 ± 0.85*	−4.1 ± 10.7	−2.5 ± 12.8*
Red (430)	8.5	−0.27 ± 0.75	−0.26 ± 0.92	−2.3 ± 8.7	−1.8 ± 10.6
Purple (764)	10.5	−0.48 ± 1.13	−0.59 ± 1.28*	−3.4 ± 9.5	−4.1 ± 10.9*
Yellow (884)	13.0	−1.03 ± 1.50	−0.92 ± 1.72*	−6.4 ± 9.5	−5.3 ± 11.1*
White (1,012)	16.5	−1.38 ± 1.92	−1.16 ± 2.14*	−6.9 ± 9.3	−5.3 ± 10.8*
Blue (923)	21.0	−1.67 ± 3.04	−1.01 ± 3.21*	−6.3 ± 10.8	−2.9 ± 12.0*
Orange (1,013)	27.0	−2.75 ± 5.10	−1.29 ± 5.54*	−7.3 ± 13.7	−1.5 ± 16.1*
Green (770) ≤143.6 cm	33.0	−4.81 ± 7.44	−3.32 ± 7.89*,†	−10.2 ± 15.3	−5.5 ± 17.3*,†
Green (142) >143.6 cm	33.0		−8.41 ± 7.71†		−17.7 ± 14.5†

Weight difference indicates the predicted weight minus the actual weight. Percent difference was calculated as follows: ([predicted weight − actual weight]/[actual weight]) × 100.
* $p < 0.05$, between 1998 and 2002A tapes.
† $p < 0.05$, between green ≤143.6 cm and green >143.6 cm for 2002A tape.

Table 7
Accuracy of the Broselow System for Predicting Actual Weight

	Accurate (Within 10%)	Inaccurate (≥10% Error*)		Inaccurate (≥20% Error†)	
		Under	Over	Under	Over
All ages					
1998 tape	60.0‡	32.7	7.3*	5.0	1.3†
2002A tape	55.3‡	31.3	13.4*	10.4	3.0†
Infants					
1998 tape	59.5‡,§	28.0	12.5*	5.1	3.3†
2002A tape	55.0‡,§	31.1	13.9*	7.5	3.5†
Toddlers					
1998 tape	64.3‡,§,‖	30.7	5.0*	5.6	0.2†
2002A tape	60.3‡,§, ‖	30.1	9.6*	6.1	1.6†
Preschool					
1998 tape	61.4‡	35.7	2.9*	8.1	0.3†
2002A tape	59.4‡,¶	32.8	7.8*	9.3	1.0†
School age					
1998 tape	57.6‡,‖	36.2	6.2*	15.5	0.7†
2002A tape	51.2‡,§,‖,¶	31.5	17.3*	15.1	4.2†

All values are expressed as percentages.
* $p < 0.05$, error ≥10% over vs. under.
† $p < 0.05$, error ≥20% over vs. under.
‡ $p < 0.05$, between 1998 and 2002A tape.
§ $p < 0.05$, infants vs. toddlers, preschool and school age.
‖ $p < 0.05$, toddlers vs. preschool and school age.
¶ $p < 0.05$, preschool vs. school age.

smaller rather than a larger size in all four age categories ($p < 0.001$).

Gender, but not race, affected the concordance of the 2002A tape. The bias (percent weight difference) was greater for male patients than for female patients (−4.8 [±13.2] vs. −3.5 [±13.4]; $p < 0.05$). When stratified by age, the difference between male patients and female patients was only noted in toddlers (−5.9 [±10.2] vs. −2.8 [±11.3]; $p < 0.05$) and preschoolers (−6.2 [±11.0] vs. −4.5 [±11.6]; $p < 0.05$). There were no gender differences noted in the PA with the Broselow color by height and the color by actual weight (moderate definition, male vs. female: 64.8% vs. 66.4%; $p = 0.123$). The bias (percent weight difference) between African Americans and non–African Americans was similar (−4.1 [±13.5] vs. −4.2 [± 13.2]; $p = 0.645$). The PA for Broselow color was also similar between races (moderate definition, African Americans vs. non–African Americans: 67.6% vs. 65.7%; $p = 0.11$).

DISCUSSION

The care and resuscitation of critically ill and injured children requires a thorough and meticulous approach. Health care professionals often rely on a rapid assessment tool, which leads to an estimate of a child's weight to precisely determine dosages of medications and to serve as a guide for choosing properly sized equipment. Care must be taken to accurately predict a patient's weight to avoid underdosing or overdosing vital medications and to precisely guide other weight-sensitive interventions.

Taking the time to place a critically ill or injured child on a scale is not practical in the majority of emergent clinical situations. Several methods have been used to predict a child's weight from a scientific guess, a parent's estimate, and an age-related resource card.

In the 1970s, Dr. Broselow developed a method of estimating a child's weight where a child's weight is estimated based on their height.[1] Since then, the Broselow system has become a common tool used in emergency departments and in out-of-hospital systems across the United States and internationally. When using a length-based tape, however, the opportunity for error is significant. Our study was designed to reassess the accuracy of the Broselow system in the face of both an epidemic of American obesity and to assess the reliability of the weight-for-length system.

The 1998 tape correctly identified the child's color zone in 66% of the children. The best concordance was in the infant group (PA, 81.3%; $\kappa = 0.66$), with the other three age groups falling below the threshold for acceptable concordance. In predicting actual weight, a European study evaluating this tape system noted a mean bias of −0.05 kg in children weighing less than 20 kg and −1.05 kg in those greater than 20 kg.[9] The mean bias in our study was worse (white zone or less, −0.16 to −1.38 kg; blue zone or greater, −1.67 to −4.81 kg). In addition, the percent of children whose predicted weight fell within 10% of their actual weight was 65% in the European study versus 60% reported in our study. The difference may be secondary to the increasing rate of obesity, especially in urban America.[14,15]

To the best of our knowledge, this study is the first independent evaluation of the 2002A tape. The overall concordance in determining color zones by height for the 1998 and 2002A tapes was similar, with a PA of 66% for both tapes. The changes made in the 2002A tape (primarily length of orange and green zones) improved the concordance for determining color zones for school-age children ($\kappa = 0.40$). The percentage of children whose weight was estimated within 10% was lower for the 2002A tape. This is in part secondary to the change in medication dosing from 1- to 2-kg increments in the 1998 tape to one weight for each color zone in the 2002A tape.

For school-age children, the concordance was better in those from the suburban cohort. These children, in contrast to the urban clinic cohort, had a lower rate of obesity, and there was a lower rate of African Americans. Because there were no differences in accuracy secondary to race, the lower concordance in urban, school-age children is probably secondary to their higher rate of obesity. Lower socioeconomic status is a significant risk factor for the development of childhood obesity and is more likely to be present in inner cities, in contrast to the affluent suburbs.[14,15] Gender differences were similar except for a slight increase in negative bias in male toddlers and preschoolers. Health care providers who use length-base systems for estimating weight need to be aware of the prevalence of obesity in their community, with higher rates reducing the accuracy of the tape.[16]

The length-based system estimated within 10% of the child's actual weight in 55.3% (2002A tape) and 60% (1998 tape) of the children. This rate is less than the 80% accuracy rate obtained from parent's estimation of their child's weight.[17] Because these predicted weights are directly utilized for dosing resuscitation medications according to PALS guidelines, the accuracy of the estimated weight also represents the accuracy of the dosing. Hence, the length-based system accurately estimated

Table 8
The Accuracy of the Broselow Tape versus the PALS Age-based Guidelines in Estimating Uncuffed Endotracheal Tubes in Children

Accurate ETT Size			Inaccurate ETT Size			
			Smaller Size by:		Larger Size by:	
Version of Tape	Percent Agreement	κ	0.5 mm (% agreement)	1.0 mm (% agreement)	0.5 mm (% agreement)	1.0 mm (% agreement)
All						
1998	71.1%	0.65	23.3%	0.7%	4.8%	0.0%
2002A	71.3%	0.66	20.9%	0.4%	7.3%	0.1%
Infants						
1998	92.4%*	0.32	7.6%	0.0%	0.1%	0.0%
2002A	92.4%*	0.32	7.6%	0.0%	0.1%	0.0%
Toddlers						
1998	69.8%*,†	0.40	27.0%	0.3%	2.9%	0.0%
2002A	67.9%*,†	0.37	29.2%	0.3%	2.7%	0.0%
Preschool						
1998	64.6%*,†,‡	§	26.6%	0.2%	8.4%	0.1%
2002A	63.6%*,†,‡	§	26.6%	0.2%	9.5%	0.1%
School age						
1998	57.6%*,‡,‖	0.35	32.5%	1.8%	8.1%	0.0%
2002A	60.2%*,†,‡,‖	0.40	24.8%	0.8%	14.1%	0.1%

* $p < 0.05$, infants vs toddlers, preschool, school age.
† $p < 0.05$, toddlers vs. preschool, school age.
‡ $p < 0.05$, preschool vs. school age.
§ Unable to calculate Kappa, only one size endotracheal tube for that age.
‖ $p < 0.05$, 1998 vs. 2002A tapes.

drug dosages in 55%–60% of the children, with the system more likely to underestimate than overestimate drug dosages. For drugs whose volume of distribution is primarily based on lean body mass, the underestimation secondary to a higher fat mass would probably have little effect on the outcome. For lipophilic drugs (i.e., midazolam,[18] amiodarone[19]), the effect of this underestimation could be magnified. In resuscitations, medications are dosed to effect. If the original dose for treating an arrhythmia or seizure was an underdosing resulting in no effect, the patient would most likely be redosed. This would lengthen the time of the resuscitation and potentially result in an adverse outcome. With the rising rate of obesity, ideal drug dosing may best be based on body composition coupled with the pharmacokinetics of the agent.[20]

The concordance between the PALS age-based guidelines and the tape's length-based recommendation for an uncuffed endotracheal tube was a PA of 71% with a κ of 0.65–0.66. The tape system was more likely to recommend a smaller size than the age-based guidelines. In the initial study based on air leak, the tape system, in contrast to the age-based formula, was more likely to select the proper size endotracheal tube (77% vs. 47%).[21,22] In two subsequent studies, one found the tape system to perform better than the age-based formula (55% vs. 41%) but more likely to select an undersized tube, whereas the other study found them to be similar.[9,23] The selection of an undersized tube may result in a significant air leak that may interfere with the ventilation and oxygenation of a poorly compliant lung (e.g., secondary to a near-drowning). The selection of a cuffed endotracheal tube (inside diameter 0.5 mm smaller than the uncuffed tube recommended by the tape) may reduce the potential for an air leak but may decrease the margin of safety and increase airway resistance.[24]

In cardiopulmonary resuscitation simulations, the use of the color-coded tape has improved the accuracy in the selection of equipment and drug dosages.[25] With an accuracy rate noted in this study ranging from 55.3% (drug dosages) to 71.3% (endotracheal tube selection), there is still the potential for improvement. The bias of the color-coded system may be enhanced by redefining the color zones so that the difference between predicted and actual weight is zero. One possible way to improve both bias and precision would be to incorporate a second measurement to assess obesity, such as midarm circumference.[26] Adding a second measurement would increase the complexity of estimating weight and may be impractical during a resuscitation. An alternative to color-coded systems would be for society to adopt a universal electronic medical record for all children that would be readily available to every provider of emergency care to children. At each well-child visit, the child's anthropometric measurement would be added to the electronic medical record. Through a drop-down window, all equipment and resuscitation medication dosages would be displayed to emergency health care providers based on those measurements. If accessed between well-child visits, the dosages and equipment sizes would be recalculated automatically based on the child's prior growth curve percentiles in relation to their age at the time of the emergency. In addition, the drug dosages could automatically be adjusted by the child's body composition, taking into account the medication's volume of distribution and lipophilic nature.

LIMITATIONS

Potential limitations of our study include a underrepresentation of African Americans in the suburban cohort and the retrospective collection of data abstracted from well-child visits in the urban clinic. In the urban clinic cohort, the weight of those subjects who fell outside the predicted color category was consistent with the patient's growth curve except for 15 subjects where transcription errors occurred at the time of their well-child visit (i.e., recording pounds as kilograms). The transcription error for this small percentage of patients was corrected in this data set so it would be consistent with their respective growth curves. Additionally, measurement of a cardboard tape with folds was difficult; securing the tape to a hard surface minimized error. Tape calibration measurements on repeat occasions were within ±1 mm. Finally, the weights for the color zones are printed on the tapes as discrete whole numbers, with 1–2 kg separating the color zones. The assignment of a color zone to children whose weight falls in those gaps can affect concordance. To address this issue, a sensitivity analysis was conducted using three separate definitions for each color zone (liberal, moderate, conservative), with the moderate definition dividing the gap in half and assigning the respective halves to the neighboring zones. In a review article, one of the authors of the Broselow system used primarily the moderate definitions of the color categories.[6] In our study, all three possibilities are presented for comparison.

The strength of our study is the use of two different cohorts coupled with consecutive enrollment. In addition, the height and weight measurements were verified prospectively by a second observer in the suburban cohort and by reviewing a subset of growth curves in the urban cohort. This study also evaluated both tapes, which are frequently used across the country.

CONCLUSIONS

We conducted an analysis of the current Broselow weight-for-length–based resuscitation system and found that the system incorrectly predicted the actual weight category in almost one third of children. This system was more likely to underestimate weight, which may in part be secondary to the rising rate of obesity in children. Caregivers need to consider this observation when using tools to estimate weight during the resuscitation of a child.

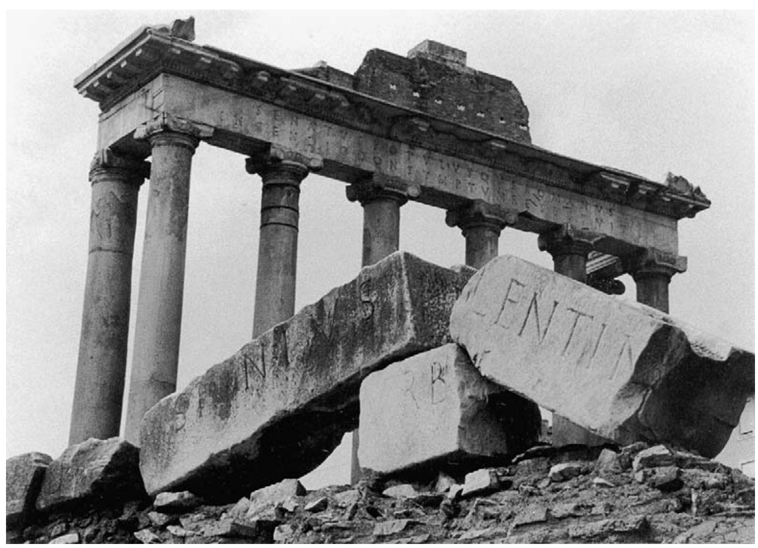
The authors thank the following for their direction, support, and assistance with this project: Jim Merlino, MD, Michael Anderson, MD, Lisa Lorenz, RN, Harry Rees, RN, Kevin Nieman, Dan Lorenz, Betty Kovach, and the IS Department at MetroHealth Medical Center.

References

1. Broselow/Hinkle Pediatric Emergency System. 1993. Vital Signs, Inc., Patent R7331I-010494.
2. Lubitz DS, Seidel JS, Chameides L, Luten R, Zaritsky A, Campbell F. A rapid method for estimating weight and resuscitation drug dosage from length in the pediatric age group. Ann Emerg Med. 1988; 17:576–81.
3. Centers for Disease Control and Prevention. Update: prevalence of overweight among children, adolescents, and adults–United States, 1988-1994. MMWR Morb Mortal Wkly Rep. 1997; 46:199–202.
4. Strauss R. Childhood obesity. Pediatr Clin North Am. 2002; 49:175–201.
5. Shah AN, Frush K, Wears R. Effects of intervention standardization system on pediatric dosing and equipment size determination: a crossover trial involving simulated resuscitation events. Arch Pediatr Adolesc Med. 2003; 157:229–36.
6. Luten R. Error and time delay in pediatric trauma resuscitation: addressing the problem with color-coded resuscitation aids. Surg Clin North Am. 2002; 82: 303–14.
7. Luten R, Wears R, Broselow J, Croskerry P, Joseph M, Frush K. Managing the unique size related issues of pediatric resuscitation: reducing cognitive load with resuscitation aids. Acad Emerg Med. 2002; 9:840–7.
8. Black K, Barnett P, Wolfe R, Young S. Are methods used to estimate weight in children accurate? Emerg Med. 2002; 14:160–5.
9. Hofer C, Ganter M, Tucci M, Klaghofer R, Zollinger A. How reliable is length-based determination of body weight and tracheal tube size in the pediatric age group? The Broselow tape reconsidered. Br J Anaesth. 2001; 88:283–5.
10. Hazinski MF, ed. Pediatric Advanced Life Support Provider Manual. Dallas, TX: American Heart Association, 2002.
11. Knoke JD. Multiple comparisons with dichotomous data. J Am Stat Assoc. 1976; 71:849–53.
12. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurements. Lancet. 1986; 1:307–10.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–74.
14. Haas JS, Lee LB, Kaplan CP, et al. The association of race, socioeconomic status, and health insurance status with the prevalence of overweight among children and adolescents. Am J Public Health. 2003; 93:2105–10.
15. Kristensen PL, Wedderkopp N, Moller NC, et al. Tracking and prevalence of cardiovascular disease risk factors across socio-economic classes: a longitudinal substudy of the European youth heart study. BMC Public Health. 2006; 6:20.
16. Theron L, Adams A, Jansen K, Robinson E. Emergency weight estimation in Pacific Island and Maori children who are large-for-age. Emerg Med Aust. 2005; 17:238–43.
17. Leffler S, Hayes M. Analysis of parental estimates of children's weights in the ED. Ann Emerg Med. 1997; 30:167–70.
18. Reilly CS, Nimmo WS. New intravenous anesthetics and neuromuscular blocking drugs. A review of their properties and clinical use. Drugs. 1987; 34:98–135.
19. Holt DW, Tucker GT, Jackson PR, et al. Amiodarone pharmacokinetics. Am Heart J. 1983; 106:840–7.
20. Duffull SD, Dooley MJ, Green B, Poole SG, Kirkpatrick C. A standard weight descriptor for dose adjustment in the obese patient. Clin Pharmacokinet. 2004; 43:1167–78.
21. Luten RC, Wears RL, Broselow J, et al. Length-based endotracheal tube and emergency equipment in pediatrics. Ann Emerg Med. 1992; 21:900–4.

22. Phipps LM, Thomas NJ, Gilmore RK, et al. Prospective assessment of guidelines for determining appropriate depth of endotracheal tube placement in children. Pediatr Crit Care Med. 2005; 6:606–8.
23. Davis D, Barbee L, Ririe D. Pediatric endotracheal tube selection: a comparison of age-based and height-based criteria. Am Assoc Nurse Anesth J. 1998; 66:299–303.
24. Ho AM-H, Aun CST, Karmakar MK. The margin of safety associated with the use of cuffed pediatric tracheal tubes. Anesthesia. 2002; 57:169–82.
25. Agarwal S, Swanson S, Murphy A, Yaeger K, Sharek P, Halamek LP. Comparing the utility of a standard pediatric resuscitation cart with a pediatric resuscitation cart based on the Broselow tape: a randomized, controlled, crossover trial involving simulated resuscitation scenarios. Pediatrics. 2005; 116:e326–33.
26. Chomtho S, Fewtrell MS, Jaffe A, Williams JE, Wells JC. Evaluation of arm anthropometry for assessing pediatric body composition: evidence from healthy and sick children. Pediatr Res. 2006; 59:860–5.

REFLECTIONS



James E. Olson, PhD
Department of Emergency Medicine
Wright State University
Dayton, Ohio

Justice Before Your Peers

While weighing the pros and cons,
Home didactics do at least spare you
From an extended stay at the Marriott.
Away forays are more fulfilling
When a voluntary main room crowd
Accepts your practiced message.
Involuntarily sequestered day travelers
Pulled from a voter registration block
Provide no audible applause at your dismissal.
Whether your artistry arches landscapes
Or monumentalizes a loathsome verdict
May sadly bypass all intellectual design.

Jonathan Singer, MD
Department of Emergency Medicine
Wright State University
Dayton, Ohio
doi: 10.1197/j.aem.2006.01.001