

# Accuracy of weight estimation by the Broselow tape is substantially improved by including a visual assessment of body habitus

Mike Wells<sup>1</sup>, Lara Goldstein<sup>1</sup> and Alison Bentley<sup>1</sup>

**BACKGROUND:** The Broselow tape (BT) has been shown to estimate weight poorly primarily because of variations in body habitus. The manufacturers have suggested that a visual assessment of habitus may be used to increase its performance. This study evaluated the ability of habitus-modified models to improve the accuracy thereof.

**METHODS:** A *post hoc* analysis of prospectively collected data from four hospitals in Johannesburg, South Africa, on a population of 1,085 children. Sixteen *a priori* models generated a modified weight estimation or drug dose based on the BT weight and a gestalt assessment of habitus.

**RESULTS:** The habitus-modified method suggested by the manufacturer did not improve the accuracy of the BT. Five dosing and four weight-estimation models were identified that markedly improved dosing and weight estimation accuracy, respectively. The best dosing model improved dosing accuracy (doses within 10% of correct dose) from 52.0 to 69.6% and reduced critical dosing errors from 16.5 to 4.3%. The best weight-estimation model improved accuracy from 59.4 to 81.9% and reduced critical errors from 11.8 to 1.9%.

**CONCLUSION:** The accuracy of the BT as a drug-dosing and weight-estimation device can be substantially improved by including an appraisal of body habitus in the methodology.

Is it inevitable that the “use of the Broselow tape may result in the under-resuscitation of children” in high-income countries and that it may dangerously overestimate weight in low- and middle-income countries such as the “South Sudan, the hungriest place on earth” (1–4)? The developers of the BT acknowledged, even in the first publication, that body habitus might need to be used to increase the accuracy of weight estimation, especially in older and heavier children (5). Their recommendation was that a child be “bumped up” a color zone if the child appears overweight or obese, to increase the accuracy of weight estimation (6,7). More recently, the manufacturer has suggested that, based on growth chart data, using this technique will ensure that more than 80% of

children will fall into the correct color zone on the 2011 edition A tape (Figure 1) (7). This has never been tested in a clinical study.

There have also been several recent studies that have shown that the BT substantially, and potentially dangerously, overestimated weight in low- and middle-income countries and in poor populations with a high prevalence of underweight children (2,8–11). The most recent edition of the BT (the 2011 edition A) was considerably modified from the previous edition to reduce underestimation of weight, especially in older children, which has magnified the overestimation of weight in these populations. The developers of the BT have not recommended a solution to this challenge, other than to suggest that emaciated children may be more tolerant of medication overdoses than children with optimum weight and nutrition (12). This viewpoint is not universally accepted, however (13).

## The Importance of Accurate Weight Estimation

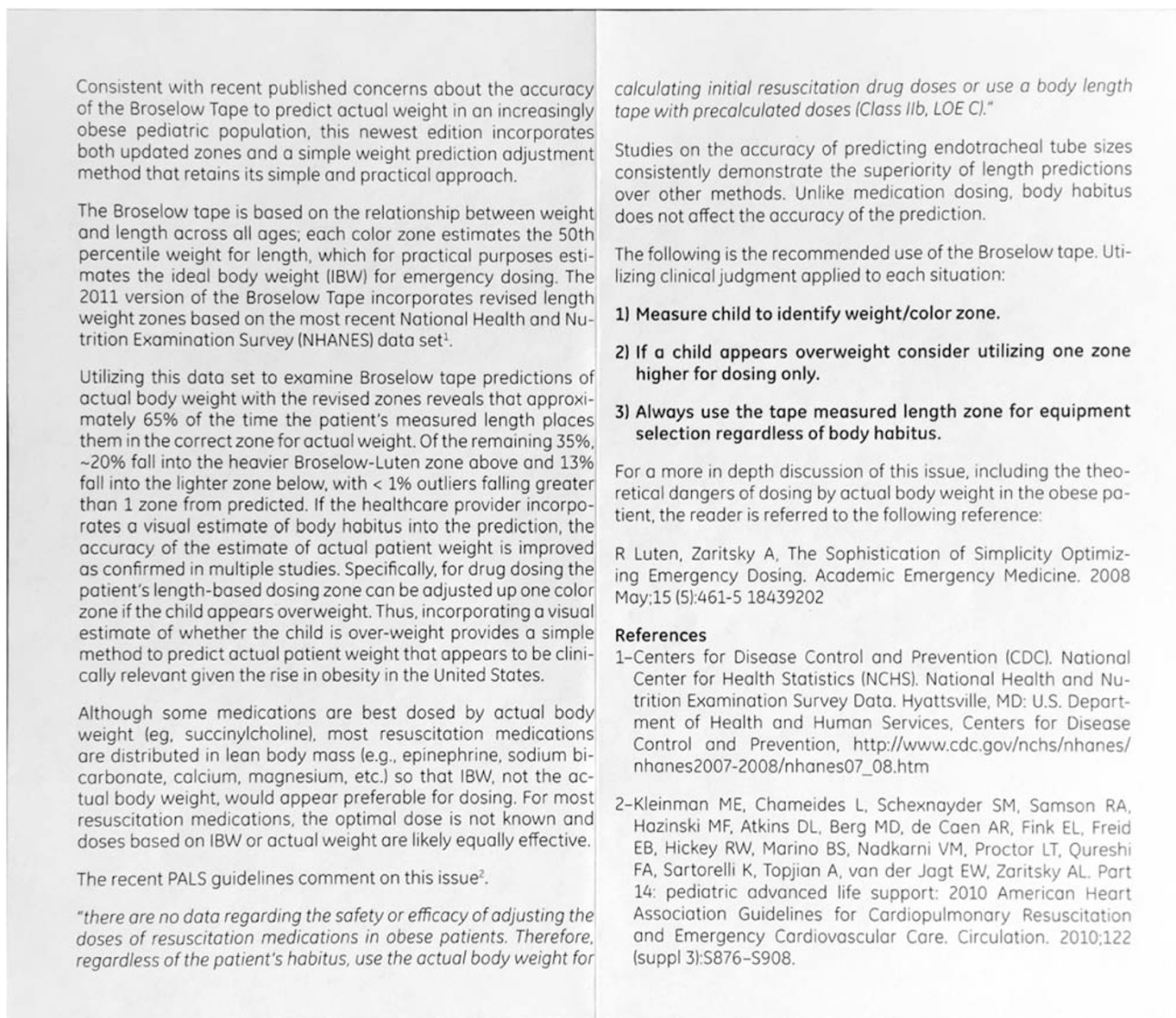
Accurate weight estimation during the management of medical emergencies in children is important to allow safe and effective administration of medications, fluids, and other weight-determined emergent therapy (14). There is sufficient evidence to show that children are at risk of sustaining medication errors during emergency care, and it is therefore mandatory that accurate and appropriate methods of weight estimation be employed (15,16). With newer dual length- and habitus-based weight-estimation methodologies, such as the Mercy method (17) and the PAWPER tape (18) having shown a much greater accuracy in weight estimation than the BT across a wide spectrum of populations, the performance of the BT must be questioned, unless the use of habitus-modified weight estimation can increase the accuracy of the tape (19,20).

## Goals of this Investigation

Our hypothesis was that the addition of a visual estimate of body habitus could significantly increase the accuracy of the BT. We aimed to test several *a priori* models of habitus-based

<sup>1</sup>Division of Emergency Medicine, Faculty of Health Sciences, University of the Witwatersrand, Parktown, Johannesburg, South Africa. Correspondence: Mike Wells (mike.wells@emergencymedicine.co.za)

Received 14 April 2017; accepted 29 August 2017; advance online publication 18 October 2017. doi:10.1038/pr.2017.222



**Figure 1.** The “package insert” from the Broselow tape 2011 edition A. This insert includes the manufacturer’s analysis of the predicted accuracy of the Broselow tape and the proposal to improve its performance, given the large number of published studies that have raised concerns about the underestimation of weight by the Broselow tape. The discussion on IBW vs. TBW is based on theoretical considerations, which have been superseded by more recent work on the subject.

modification of BT weight estimation and of habitus-based modification of BT drug doses, including the model recommended by the manufacturer.

## METHODS

This was an analysis of pooled data from three previous studies evaluating various weight estimation systems, including the BT (11,18,21). Permission to conduct the original studies as well as the secondary data analysis was obtained from the Human Research Ethics Committee of the University of the Witwatersrand. Written informed consent was obtained from parents and written assent was obtained from children over the age of 7 years.

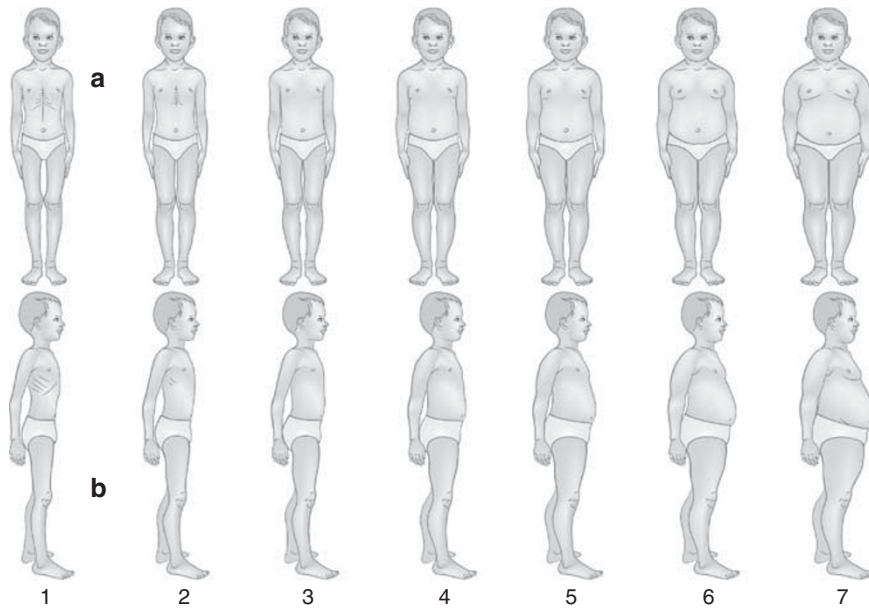
## Study Design and Setting

Each study was a prospective, observational, cross-sectional study from four hospitals in Johannesburg, South Africa. Two of the

hospitals serve a community of mostly low socioeconomic status and two serve a middle-class community. Study 1 enrolled 453 children from September 2008 to October 2008 in two centers. Study 2 enrolled 332 children from July 2014 to December 2014 in one center and Study 3 enrolled 300 children from August 2014 to January 2015 in a single center. The aims of these studies were to evaluate the weight estimation accuracy of the PAWPER tape, the Mercy method, and the original BT methodology.

## Selection of Participants

These data were obtained from convenience samples of children not requiring emergency medical care. The inclusion criteria were children aged 1 month to 12 years (Studies 1 and 3) or 1 month to 16 years (Study 2). The exclusion criteria were children with congenitally abnormal stature (e.g., dwarfism) and children whose length could not be assessed (e.g., because of contractures).



**Figure 2.** Figural representations of children in the seven habitus score (HS) categories. These are representations of a 6-year-old child (**a** shows the frontal views and **b** shows the side views; the numbers show the HSs). From left to right: skinny (1), slightly thin (2), optimum (3), overweight (4), obese (5 and 6), and severely obese (7). A HS of 1 was designed to roughly correspond to a weight at the 5th centile of weight-for-length, a score of 2 with the 25th centile, a score of 3 (average) with the 50th centile, a score of 4 with the 75th centile, and a score of 5 with the 95th centile and HSs of 6 and 7 at the upper ends of the centiles. The use of reference images for real-time habitus assessment may allow standardization as different populations may have a different perception of the “average” child.

## Methods and Measurements

**Study procedure.** A standardized procedure was followed in each study:

- Each child’s supine length was measured with a standard measuring tape.
- The BT (2007 edition A for Study 1 and 2011 edition A for Studies 2 and 3) was used to generate an estimate of weight, according to the instructions on the tape.
- A visual inspection of the child was performed to assign a numerical quantification of body habitus according to previously described methods (Figure 2) (18).
- The child’s actual weight was then measured on a scale and recorded to the nearest 0.1 kg (Tanita SC-240 Body Composition Monitor—Class III device accurate to  $\pm 0.1$  kg).

**Data extraction for the post hoc analysis.** Records from each of the contributing studies were pooled for the *post hoc* analysis; complete data were available for every child. The data that were used from each of the 1,085 children in this analysis included the following: age, sex, supine length, BT weight, BT color zone, habitus score (HS), and actual measured weight. Z-scores of BMI-for-age were calculated for each child (using the World Health Organization BMI-for-age growth data for children under the age of 2 years and growth data from the Centers for Disease Control (2000) charts for children over the age of 2 years).

**Model development.** The models to be tested (Table 1) were developed *a priori*, based on the findings of previous studies (22–25). The models were constrained to be simple enough to be used manually or with a simple mobile phone app. The model proposed by the manufacturers of the BT was also included.

## Analysis

Following four primary statistical measures were used to assess performance:

- Mean percentage error was calculated for each model as a quantification of the overall estimation bias.
- The 95% limits of agreement of the mean percentage error (PELOA) were calculated to represent a quantification of the estimation precision.
- The percentage of drug dose recommendations or weight estimations that fell within 10% (PW10) and 20% (PW20) of true drug dose or measured weight was calculated for each model to represent a quantification of overall accuracy.
- The absolute percentage deviation from the recommended dose range, using adrenaline as an exemplar, was used to assess drug-dosing accuracy after the method of Shah *et al.* (26).

Subgroup analyses were performed using the same outcome measures and statistical analyses for children in the following three weight categories: <10, 10–25, and >25 kg in underweight, normal weight, and overweight/obese children.

The data were analyzed to evaluate differences between the models, the original BT data, and the gold standard using a modified Bland–Altman system, comparison of the mean percentage error and PW10 analysis with parametric (paired *t*) and categorical (McNemar) tests. A  $P < 0.05$  was regarded as significant for all statistical tests.

## Outcomes

The primary outcome measure was the performance of each of the modified Broselow models with respect to drug-dose correctness, when compared with a dose based on measured weight. For this study, a drug-dose deviation of >10% of the calculated true dose was considered to be an incorrect dose and a dose deviation of >20% of the actual dose was considered to be a critical dose error (27,28). An improvement on the current BT accuracy of >10% was regarded as significant.

The secondary outcome measure was the performance of each of the modified-Broselow weight-estimation models when compared with the unmodified Broselow weight as well as to actual measured

**Table 1.** Both dosing (prefixed by “D”) and weight-estimation (prefixed by “W”) models were conceptualized *a priori* to test the principle of habitus-modified functioning of the Broselow tape

| Model                           | Description of model   | Outcome                        |
|---------------------------------|--|--------------------------------|
| <i>Drug-dosing models</i>       |  |                                |
| D1                              | Increase color zone only for HS > 3  | No improvement in outcome      |
| D2                              | Increase color zone only for HS > 4  | No improvement in outcome      |
| D3                              | Increase color zone for HS > 3 and decrease color zone for HS < 3  | No improvement in outcome      |
| D4                              | Increase color zone for HS > 4 and decrease color zone for HS < 3  | Final ranking: 5 <sup>th</sup> |
| D5                              | Increase color zone for HS > 4, increase GREEN by 10%, and decrease color zone for HS < 3  | Final ranking: 4 <sup>th</sup> |
| D6                              | Increase color zone by one for HS 4 and two for HS ≥ 5 and decrease color zone for HS < 3  | No improvement in outcome      |
| D7                              | Increase dose by 10% for HS > 3 and decrease dose by 10% for HS < 3  | Final ranking: 3 <sup>rd</sup> |
| D8                              | Increase dose by 10% for HS 4 and 20% for HS ≥ 5 and decrease dose by 10% for HS < 3   | Final ranking: 2 <sup>nd</sup> |
| D9                              | Increase dose by 10% for HS 4 and 20% for HS ≥ 5 and decrease dose by 10% for HS 2 and by 20% for HS 1   | Final ranking: 1 <sup>st</sup> |
| <i>Weight-estimation models</i> |  |                                |
| W1                              | For HS > 3 increase to the equivalent weight one color zone up (beginning, middle, or end of zone)   | Rules ambiguous and complex    |
| W2                              | For HS > 3 increase to the equivalent weight one color zone up, for HS < 3 decrease to the equivalent weight one zone down (beginning, middle, or end of zone)   | Rules ambiguous and complex    |
| W3                              | For HS 4 and HS 5 increase to the equivalent weight one color zone up, for HS 6 and HS 7 increase to the equivalent weight one color zone up, for HS < 3 decrease to the equivalent weight one zone down (beginning, middle, or end of zone) | Rules too complex              |
| W4                              | For HS > 3 increase weight estimation by 10% and for HS < 3 decrease weight estimation by 10%  | Final ranking: 2 <sup>nd</sup> |
| W5                              | For HS 4 increase the weight estimation by 10%, for HS ≥ 5 increase the weight estimation by 20%, for HS 2 decrease the weight estimation by 10%, and for HS 1 decrease by 20%   | Final ranking: 1 <sup>st</sup> |
| W6                              | For HS > 3 increase by one weight zone and for HS < 3 decrease by one weight zone  | Final ranking: 4 <sup>th</sup> |
| W7                              | For HS 4 increase by one weight zone, for HS ≥ 5 increase by two weight zones, for HS 2 decrease by one weight zone, and for HS 1 decrease by two weight zones   | Final ranking: 3 <sup>rd</sup> |

D1 and D2 models are based on the manufacturer’s recommendations. Models were of three basic constructs: (i) either based entirely on the Broselow tape, with shifts up or down color- or weight zones depending on habitus, (ii) with modifications of Broselow tape estimated dose or weight up or down in multiples of 10% depending on habitus, (iii) or a blend of both methodologies. The blended models were used for children in the green color zone who could not have an upward “bump” in zone, but rather had a 10% increase in dose. Increases in weight estimation were not limited to the maximum of the Broselow tape (36 kg) but were increased as per the model rules (e.g. an obese child falling into the 36kg green zone could have their weight “bumped up” two weight-zones to a maximum of 40kg). Some models were rejected before the testing phase because of ambiguities of the phrasing and because of complexities that would not have been feasible to be used during the stress of emergency medical care. The value of 10% for modifications in the nominal models was selected primarily for its simplicity in calculations, but also because previous studies from diverse populations have shown that adjustments of 8– 12% (up in fat populations and down in thin populations) would have been appropriate.

weight. A PW10 of > 70% and a PW20 > 95% were considered to be an acceptable outcome of the model, as has been previously proposed (29). This is the level of accuracy consistently achieved by other length-based, habitus-modified weight-estimation systems.

**RESULTS**

**Characteristics of Study Participants**

The basic demographic information is shown in Table 2. The study population was reasonably well distributed across the age- and weight spectrum of children for whom the BT can be used, although infants were under-represented. Just over 5% of the total sample (59 children) were too tall for the tape (> 145 cm). Only 15 (25.4%) of these children were within 10% of the maximum BT weight and only 24 (40.7%) were

close to an “adult” weight of 50 kg. The remaining 20 children (33.9%) had a substantially lower weight than “adult” weight, but significantly higher weight than the maximum possible BT weight. No children were excluded from the studies because of an inability to use the tapes (e.g., because of contractures).

The distribution of BMI-for-age Z-scores, and the relationship between Z-scores and HS assignment for the study population, is shown in Figure 3. The evaluation of inter-rater agreement for the assessment of body habitus in a subsample of 134 children showed a percentage agreement of 82.1%, with a Cohen’s kappa of 0.74 (0.64, 0.83) and a linear-weighted kappa of 0.81 (0.74, 0.89), which indicated a good to very

**Table 2.** Basic demographic information from the study population which was relevant to the study design

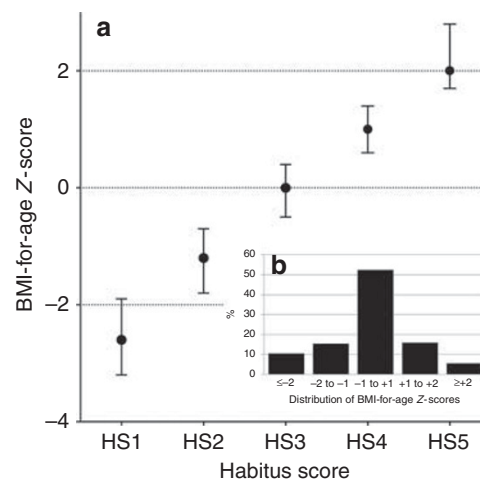
|                             | All              | ≤ 10 kg         | 10–25 kg         | > 25 kg          | HS1              | HS2              | HS3              | HS4              | HS5  | HS6              | HS7      |
|-----------------------------|------------------|-----------------|------------------|------------------|------------------|------------------|------------------|------------------|--|------------------|----------|
| Number (%)                  | 1,085            | 125 (11.5)      | 669 (61.7)       | 291(26.8)        | 63 (5.8)         | 241 (22.2)       | 460 (42.3)       | 245 (22.6)       | 61 (5.6)   | 12 (1.1)         | 3 (0.3)  |
| Male (%)                    | 561 (51.7)       | 64 (51.2)       | 353 (52.7)       | 144 (49.4)       | 37 (60.7)        | 112 (46.4)       | 255 (55.4)       | 116 (47.3)       | 35 (57.3)  | 5 (41.7)         | 1 (33.3) |
| Age (median)                | 4.9              | 0.8             | 4.0              | 9.3              | 6.3              | 5.3              | 4.2              | 4.6              | 7.0  | 10.4             | 7.7      |
| Weight (median)             | 18.2             | 8.2             | 16.3             | 32.1             | 17.8             | 17.2             | 17.0             | 20.5             | 32.5   | 51.3             | 51.2     |
| BMI Z-score (median)        | 0.0              | -0.3            | -0.3             | 0.5              | -2.6             | -1.2             | 0                | 1.0              | 2.0  | 2.2              | 2.7      |
| Gray <i>n</i> (%)           | 14 (1.3)         | 14              | 0                | 0                | 0                | 2 (14.3)         | 10 (71.4)        | 2 (14.3)         | 0  | 0                | 0        |
| Pink <i>n</i> (%)           | 25 (2.3)         | 25              | 0                | 0                | 1 (4.0)          | 2 (8.0)          | 15 (60.0)        | 5 (20.0)         | 2 (8.0)  | 0                | 0        |
| Red <i>n</i> (%)            | 64 (5.9)         | 57 (89.1)       | 7 (10.9)         | 0                | 5 (7.8)          | 12 (18.8)        | 29 (45.3)        | 17 (26.6)        | 1 (2.6)  | 0                | 0        |
| Purple <i>n</i> (%)         | 99 (9.1)         | 26 (26.3)       | 73 (73.7)        | 0                | 8 (8.1)          | 18 (18.2)        | 43 (43.4)        | 27 (27.3)        | 3 (3.0)  | 0                | 0        |
| Yellow <i>n</i> (%)         | 155 (14.2)       | 3 (1.9)         | 152 (98.1)       | 0                | 2 (1.3)          | 32 (20.6)        | 67 (43.2)        | 45 (29.0)        | 7 (4.5)  | 2 (1.3)          | 0        |
| White <i>n</i> (%)          | 199 (18.3)       | 0               | 199              | 0                | 8 (4.0)          | 49 (24.6)        | 110 (55.3)       | 28 (14.1)        | 4 (2.0)  | 0                | 0        |
| Blue <i>n</i> (%)           | 187 (17.2)       | 0               | 175 (93.6)       | 12 (6.4)         | 13 (7.0)         | 53 (28.3)        | 82 (43.9)        | 31 (16.6)        | 8 (4.3)  | 0                | 0        |
| Orange <i>n</i> (%)         | 147 (13.5)       | 0               | 58 (39.5)        | 89 (60.5)        | 9 (6.1)          | 35 (23.8)        | 43 (29.3)        | 42 (28.6)        | 17 (11.6)  | 0                | 1 (0.7)  |
| Green <i>n</i> (%)          | 195 (18.0)       | 0               | 5 (2.6)          | 190 (97.4)       | 17 (8.7)         | 38 (19.5)        | 61 (31.3)        | 48 (24.6)        | 19 (9.7)   | 10 (5.1)         | 2 (1.0)  |
|                             | All              | ≤ 10 kg         | 10–25 kg         | > 25 kg          | HS1              | HS2              | HS3              | HS4              | HS ≥ 5   | Length > 145 cm  |          |
| Number (%)                  | 1,026            | 125 (12.2)      | 669 (65.1)       | 232 (22.6)       | 61 (5.8)         | 231 (22.5)       | 441 (42.9)       | 232 (22.6)       | 61 (6.0)   | 59 (5.4)         |          |
| Male (%)                    | 530 (51.7)       | 64 (51.2)       | 353 (52.7)       | 113 (48.7)       | 36 (59.0)        | 108 (46.7)       | 244 (55.3)       | 107 (46.1)       | 35 (57.4)  | 31 (52.5)        |          |
| Age (median) IQR            | 4.5, 2.1, 7.2    | 0.8, 0.5, 1.2   | 4.0, 2.5, 5.8    | 9.0, 7.7, 10.0   | 6.0, 3.9, 9.0    | 5.0, 2.9, 7.3    | 4.0, 2.0, 6.5    | 4.2, 1.9, 7.2    | 6.8, 3.1, 8.8  | 12.0, 11.3, 13.6 |          |
| Weight (median) IQR         | 17.5, 12.7, 24.1 | 8.2, 7.2, 9.1   | 16.3, 13.4, 20.0 | 30.6, 27.5, 34.4 | 17.0, 12.6, 23.1 | 17.0, 12.0, 21.8 | 16.4, 12.4, 21.7 | 19.5, 13.2, 29.1 | 31.5, 20.0, 40.6   | 45.2, 40.1, 56.4 |          |
| BMI Z-score (median) IQR    | 0.0, -1.0, 0.8   | -0.3, -1.2, 0.6 | -0.3, -1.2, 0.6  | 0.6, -0.2, 1.3   | -2.6, -3.2, -1.9 | -1.2, -1.8, -0.7 | 0, -0.5, 0.4     | 1.0, 0.6, 1.4    | 2.0, 1.7, 2.8  | 0.4, -0.3, 1.3   |          |
| ≤ 10 kg                     | 125 (12.2)       | 8               | —                | —                | 14 (23.0)        | 27 (11.7)        | 62 (14.1)        | 20 (8.6)         | 2 (3.3)  | —                |          |
| 10–25 kg                    | 669 (65.1)       | —               | —                | —                | 36 (70.6)        | 174 (75.3)       | 306 (69.4)       | 132 (56.9)       | 29 (4.8)   | —                |          |
| > 25 kg                     | 232 (22.6)       | —               | —                | —                | 11 (18.0)        | 30 (13.0)        | 73 (16.6)        | 80 (34.5)        | 30 (49.2)  | —                |          |
| Length > 145cm <i>n</i> (%) | 59               | —               | —                | —                | 2 (3.4)          | 10 (16.9)        | 19 (32.2)        | 13 (22.0)        | 15 (25.4)  | —                |          |
|                             | All              | ≤10 kg          | 10–25 kg         | > 25 kg          | < 1 Year         | 1–5 Years        | > 5 Years        | Length > 145 cm  | "Too tall" lengths:<br>145–150 cm, 30.5%<br>150–155 cm, 23.7%<br>> 155 cm. 45.8% |                  |          |
| Overweight                  | 11.6             | 5.6             | 8.5              | 23.7             | 8.2              | 10.7             | 13.3             | 23.7             |  |                  |          |

Table 2 Continued

|                     | All  | ≤10 kg | 10–25 kg | > 25 kg | < 1 Year | 1–5 Years | > 5 Years | Length > 145 cm | “Too tall” lengths:<br>145–150 cm, 30.5%<br>150–155 cm, 23.7%<br>> 155 cm, 45.8%                                  |
|---------------------|------|--------|----------|---------|----------|-----------|-----------|-----------------|---|
| Obese               | 5.7  | 3.2    | 4.3      | 10.8    | 4.1      | 6.0       | 5.7       | 11.9            | “Too tall” weights:<br>30–36 kg, 11.9%<br>36–40 kg, 13.6%<br>40–45 kg, 22.0%<br>45–50 kg, 11.9%<br>> 50 kg, 40.7% |
| Overweight or obese | 17.3 | 8.8    | 12.9     | 34.5    | 12.2     | 16.6      | 19.0      | 35.6            |   |
| Underweight         | 8.4  | 12.0   | 9.6      | 3.0     | 9.2      | 6.0       | 10.7      | 3.4             |   |

HS, habitus score; IQR, interquartile range.

The upper panel shows the distribution of children within the weight and HS categories as well as in the Broselow color zones. The information on the 59 children “too tall for the tape” who were excluded from the primary analysis is included. The lower panel shows the relationship between weight categories and HS categories. Children from HS 5, HS 6, and HS 7 were pooled for all analyses because of the relatively small number of children with severe obesity.



**Figure 3.** The relationship between HS and BMI-for-age Z-scores is illustrated in **a**. The median and interquartile ranges for Z-scores are plotted for each HS. The Spearman’s correlation co-efficient was 0.81 ( $P < 0.001$ ) for the relationship and each HS category was significantly different to the next (Mann-Whitney test,  $P < 0.0001$ ). The insert (**b**) is a histogram of the distribution of Z-scores within the population. According to the World Health Organization criteria, a Z-score of  $< -2$  is generally considered to identify a significantly underweight child, a Z-score of  $> +1$  indicates an overweight child, a Z-score of  $> +2$  indicates an obese child, and a Z-score between  $-1$  and  $+1$  indicates “optimum” habitus.

good strength of agreement (30). No rating differed by more than one HS category for any child.

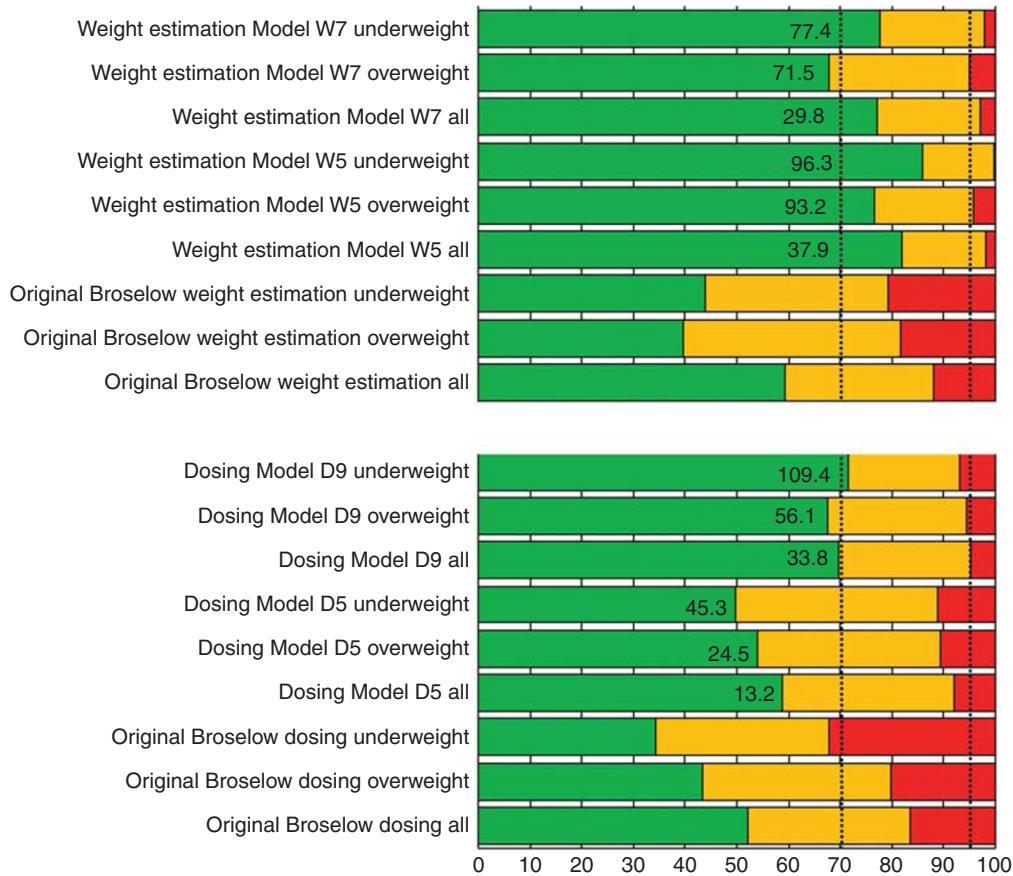
**The Performance of the Unmodified BT**

The prediction by the manufacturer of the correctness of placement of children into color zones with the unmodified BT (Figure 1) was not borne out by the findings of this study. Fewer children had a correct color-zone placement (57.0% vs. 65% predicted), slightly more children had an underestimation of weight (23.4% vs. 20% predicted), the same number of children had weight overestimated as predicted (13.2% vs. 13% predicted) but there were many more children who were more than one color-zone off (6.4% vs. <1% predicted).

The performance of the unmodified BT in terms of accuracy of drug doses and weight estimation is shown in Figure 4 and Supplementary Table S1 online. Drug doses were correct in only 52.0% of the sample, with more than 16% of critical errors, mostly doses above the recommended range. This was consistent across the weight category subgroups. Weight estimation was slightly more accurate, with a PW10 of nearly 60%, but 11% of children still received a critically erroneous weight estimation. Weight estimation accuracy fell off sharply in older and heavier children.

**The Performance of Habitus-Modified Dosing Models**

Models D1, D2, D3, and D6 worsened or failed to improve the accuracy of drug dosing and were rejected (Supplementary Table S1). The remainder of the models provided a significantly improved drug-dose accuracy with an overall small-to-medium effect size. The 10%-based models were



**Figure 4.** Weight estimation and drug-dosing accuracy of the habitus-modified models for the entire sample, for underweight children, and for overweight and obese children. The green portion of the bars indicate the percentage of weight estimations or drug doses within 10% of the correct weight or dose. The green bars plus the orange bars represent those estimates within 20% of the correct weight or dose. The red portion of the bars indicate critical errors in the estimates (error > 20%). The number in the bar indicates the relative percentage improvement of the model when compared with the performance of the unmodified Broselow tape (overall and by subgroup).

more accurate than those based on color-zone “bumping” methods, with models D5 and D9 being the best of the color-zone and 10%-based methods, respectively. The improvement of drug-dosing accuracy was substantial and the effect sizes were largest in the subgroups of heavier, thinner, and fatter children.

**The Performance of Habitus-Modified Weight-Estimation Models**

All the evaluated models (W4 to W7) showed a significant, considerable improvement in weight-estimation accuracy (medium effect size), with the 10%-based methods again outperforming the weight-zone-based models. The subgroup analysis of the W5 and W7 models (the best of each method) showed a uniform improvement across the weight spectrum, but very large effect size increases in accuracy in children of high and low weight-for-length (thin and fat).

All analyses were performed with and without the inclusion of the 59 children > 145 cm length. This made little overall difference to the model performances. Analysis with the

substitution of BT 2011 edition A for the BT 2007 edition B tape also showed no significant difference overall.

**DISCUSSION**

Only 42.9% of this study population was assessed as of “normal”, or average, habitus, which meant that nearly 60% of children required a modification of weight estimation to receive an appropriate drug dose. This demonstrates the importance of and need for habitus-based modification methodologies to ensure safe drug dosing. Although there was a reasonably high prevalence of both underweight and obese children in this study population, other populations may well have a higher prevalence of either obese or underweight children (or both), with a consequent greater need for a habitus-modified system (22,31).

**The Outcomes of the Habitus-Modified Drug-Dosing Models**

The evaluation of the habitus-modified dosing model recommended by the manufacturer not only showed no improvement, but, in fact, a slight worsening of drug-dosing accuracy. Furthermore, previous studies have shown that the

BT is at risk of both underestimation as well as overestimation of weight in different populations; therefore, any habitus-based model *must* be able to accommodate a wide spectrum of habitus types and not only apply to obese children (1,2).

Unfortunately, there are few previous studies comparing the accuracy of the actual BT drug doses against those calculated from measured weight. In most previous reports, the BT dose is assumed to be correct if the correct color zone is selected (26). This is clearly incorrect as shown in this study, with the original BT providing a “correct” dose in only half the study population and a critical dosing error in nearly 17% of the population.

The improvement in performance of the “bumping” color zone drug-dosing models was modest when compared with the unmodified tape performance, although they did provide a significant reduction in critical dosing errors. This is because the BT provides only one dose per color zone and not a dose for each weight zone (except for the gray zone). Although this allows simplicity, it increases the prevalence of drug-dosing errors. Any model based on color zones could, therefore, provide only very coarse adjustments of dose.

On the other hand, the performance of the 10%-based models, which modified the dose itself rather than the color zone, was more impressive with a sizable significant increase in accuracy. These models achieved a critical dosing error rate of less than 5%. This approach permitted a degree of fine-tuning not possible with the color-zone-bumping method.

The accuracy of the models was consistently reproduced in the subgroup analysis between weight and habitus categories, with very low critical error rates. The exception was with the D5 model in underweight children ( $HS < 3$ ) where drug doses were excessively reduced and 11% of these children would have been underdosed. Re-analysis of the data with a less stringent reduction (only for HS 1) did not improve the performance.

### The Outcomes of the Habitus-Modified Weight-Estimation Models

The relative accuracy of weight estimation was greater for both the unmodified BT as well as for both forms of the habitus-modified weight estimation models when compared with the drug-dosing accuracy. The BT has been extensively studied as a weight-estimation device, with the two best performances ever recorded originating from studies in Iran and in New Zealand with a PW10 of 72.5% and 73.4%, respectively (32,33). The two worst performances were in populations from India and South Sudan with a PW10 of 28.0% and 15.1%, respectively (2,10). The performance of the unmodified BT in this study was about average, compared with previous studies; however, many children were at risk of critical dosing errors. This accuracy is well below what the newer dual length- and habitus-based methods have achieved—a PW10 consistently greater than 70% and PW20 greater than 95%.

The performance of the weight-zone “bumping” models was substantially better than the color-zone “bumping”

models because of the higher number and closer spacing of the weight zones. The 10%-based models still performed statistically significantly better than the “bumping” models, however, at least partly as a maximum weight of more than 36 kg could be achieved (up to 40 kg for overweight, green children). The accuracy achieved by the modified weight-estimation models was exceptional and rivalled the results from other dual length- and habitus-based systems. The best models had fewer than 3% critical errors. This accuracy was remarkably well maintained throughout the weight subgroups evaluated, as well as in thin, average, and overweight children.

Given the evidence of the accuracy of these outcomes, and the simplicity of the “bumping” models (if not the 10%-based models), the stimulus to consider evaluating them further and using them in clinical practice is compelling.

### The 10%-Based vs. the Zone-Based Models

The accuracy of the 10%-based models was significantly superior to that of the color- and weight-zone “bumping” methods, especially in the dosing models. The disadvantage of this methodology is that it requires some mental arithmetic and, although it may appear to be simple to perform these calculations, there is good evidence to show that even simple arithmetic is fraught with error during stressful situations (34–36). This needs to be tested using the current models in simulated pediatric emergencies. The choice of which technique to use may depend on the nature of the emergency, the time pressure, and the balance between accuracy and ease-of-use. Both systems would require appropriate training to minimize errors (34).

### Appropriate Targets for Drug-Dosing and Weight-Estimation Accuracy

Some authors argue that an overly stringent accuracy target for weight-estimation and drug-dosing are not justified, given that many of the drug-dose guidelines are not well established by high-quality evidence (12,19). Given the absolute imperative to minimize medication errors, it is a good medical practice to eliminate as many errors as is reasonably achievable, and it is certainly not a good medical practice to continue to use a proven poor weight-estimation (and therefore drug-dosing) system (37). The use of habitus-modified models to improve the accuracy of the BT may be one potential solution to this root-cause error.

### Children too Tall for the Tape (> 145 cm)

There are problems with using the Broselow tape in children > 145 cm length, as shown previously (38). The manufacturer’s recommendation for the BT is to use an adult dose for children > 36 kg, but only 40% of these children could have been considered to have close to an “adult weight” of > 50 kg. This means that substantial drug errors would have resulted for 60% of the children by following the BT recommendations. This recommendation needs to be revised.



**How Generalizable are the Findings of this Study?**

There are two components to this important question. The first is whether the models would be as effective in populations that are different to this study population. The fact that the best models were consistently accurate in both the underweight and overweight subgroups supports the argument that the models will probably be effective in populations with a greater proportion of underweight or overweight children. Furthermore, the study sample contained enough normal weight, underweight, and overweight children to adequately test the functioning of the original BT and the proposed models.

The second question of generalizability is whether the accurate assessment of habitus can be reproduced by other non-expert users. In this study, the strong association between HS and BMI-for-age confirmed that visual inspection was successfully used to categorize habitus. Visual assessment of habitus has been shown to be accurate in previous studies (11,18,20,39), but has also been shown to be less successful without appropriate training (40,41). There are other more objective methods of assessing habitus that may prove useful with the models described in this study, such as the use of reference-image-assisted habitus assessment, or anthropometric predictors of habitus. Figural reference image-based habitus assessment has previously been established as feasible and accurate and can be used to provide a less subjective method of rapidly assessing habitus (18,39,42). Similarly, anthropometric measurements, such as mid-arm circumference, have been well established as accurate predictors of habitus and could be included in a model (17,43). The use of waist circumference has also recently also been shown to significantly improve the accuracy of weight prediction by the BT in obese children (20). Although anthropometric methods are more objective than anthroposcopic methods, they may be technically difficult to obtain during emergency medical care. This would, however, be a further interesting avenue of research as the best method of assessing habitus is yet to be established.

With respect to the habitus assessment in the different models, there was no requirement for subtle distinctions between habitus categories in the “bumping” drug-dosing models, with color zones increased only for obese children (HS 5 and above) and decreased for underweight children (HS 1 and 2). This means that this model should be very reproducible, given that the extremes of habitus are easy to identify, even for inexperienced users. The other models required a bit more refinement in differentiating the overweight (HS 4) from the obese (HS  $\geq$  5) child and the thin (HS 2) from the emaciated (HS 1) child. The evidence from this study was that this was achieved successfully, however.

**Limitations**

The development of models was limited by the architecture and design of the BT, with its inherent characteristics (one drug dose value per color zone), the need to limit models to

(relatively) cognitively neutral methods, and the uncertainty of acceptable end points (what degree of error is tolerable).

Although accurate weight estimation is important, a weight-estimation system should not add to the cognitive burden experienced by the treating team as this may create more opportunities for error (44,45). The models used in this study improved accuracy significantly, and were simple in design, but will need to be tested further to establish their cognitive impact and vulnerability to error.

**CONCLUSIONS**

The incorporation of a visual assessment of body habitus into a relatively simple model can greatly improve the accuracy of both weight estimation and drug dosing. The 10%-based methods were more accurate than the color-zone “bumping” methods but both were significantly more accurate than the unmodified BT.

Given how much more accurately the weight-estimation models improved the BT’s performance, when compared with the drug-dose models, it might be appropriate to preferentially use the habitus-modified weight estimation as the basis for dose determination.

Additional research will be required to assess the cognitive impact and risk for errors when these models are used during simulated management of medical emergencies and thus to establish their role in clinical practice. The use of mobile phone apps or computerized system may prove invaluable to reduce cognitive load as systems unavoidable become more complex.

Disclosure: We declare that we have no financial conflict of interest. Prof. Wells is the developer of the PAWPER tape, but derives no financial or commercial benefit from the tape.

**REFERENCES**

1. Nieman CT, Manacci CF, Super DM, Mancuso C, Fallon WF Jr. Use of the Broselow tape may result in the underresuscitation of children. *Acad Emerg Med* 2006;13:1011–9.
2. Clark MC, Lewis RJ, Fleischman RJ, Ogunniyi AA, Patel DS, Donaldson RI. Accuracy of the Broselow Tape in South Sudan, “The Hungriest Place on Earth”. *Acad Emerg Med* 2016;23:21–8.
3. Waseem M, Chen J, Leber M, Giambrone AE, Gerber LM. A reexamination of the accuracy of the Broselow Tape as an instrument for weight estimation. *Pediatr Emerg Care* 2017. (doi:10.1097/PEC.0000000000000982; e-pub ahead of print).
4. Young KD, Korotzer NC. Weight estimation methods in children: a systematic review. *Ann Emerg Med* 2016;68:441–51 e10.
5. Lubitz DS, Seidel JS, Chameides L, Luten RC, Zaritsky AL, Campbell FW. A rapid method for estimating weight and resuscitation drug dosages from length in the pediatric age group. *Ann Emergency Med* 1988;17: 576–81.
6. Deboer S, Seaver M, Broselow J. Color coding to reduce errors. *Am J Nurs* 2005;105:68–71.
7. Broselow tape 2011 edition A package insert. Armstrong Medical Inc. Lincolnshire, IL, USA. 2011.
8. House DR, Ngetich E, Vreeman RC, Rusyniak DE. Estimating the weight of children in Kenya: do the Broselow tape and age-based formulas measure up? *Ann Emergency Med* 2013;61:1–8.
9. Dicko A, Alhousseini ML, Sidibé B, Traoré M, Abdel-Rahman SM. Evaluation of the Mercy weight estimation method in Ouelessebougu, Mali. *BMC Public Health* 2014;14:270.

10. Batmanabane G, Kumar Jena P, Dikshit R, Abdel-Rahman SM. Using the Mercy Method for weight estimation in Indian children. *Glob Pediatric Health* 2015;1-7.
11. Georgoulas V, Wells M. The PAWPER tape and the Mercy Method outperform other methods of weight estimation in children in South Africa. *South African Med J* 2016;106:933-9.
12. Luten R, Zaritsky A. The sophistication of simplicity... optimizing emergency dosing. *Acad Emerg Med* 2008;15:461-5.
13. Wells M, Kramer E. Optimizing emergency drug dosing in children. *Acad Emerg Med* 2008;15:1325.
14. Wells M, Goldstein L, Botha M. *Emergency Drug Dosing in Children*. Edinburgh: Churchill Livingstone Elsevier, 2012.
15. Kozer E, Scolnik D, Macpherson A, Keays T, Shi K, Luk T, et al. Variables associated with medication errors in paediatric emergency medicine. *Pediatrics* 2002;110:737-42.
16. Kaufmann J, Laschat M, Wappler F. Medication errors in pediatric emergencies: a systematic analysis. *Deutsches Ärzteblatt Int* 2012;109:609-16.
17. Abdel-Rahman SM, Ridge AL. An improved pediatric weight estimation strategy. *Open Med Devices J* 2012;4:87-97.
18. Wells M, Coovadia A, Kramer E, Goldstein L. The PAWPER tape: a new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. *Resuscitation* 2013;84:227-32.
19. Milne K, Yasin A, Knight J, Noel D, Lubell R, Filler G. Ontario children have outgrown the Broselow tape. *Can J Emergency Med* 2012;14:25-30.
20. Tanner D, Negaard A, Huang R, Evans N, Hennes H. A prospective evaluation of the accuracy of weight estimation using the Broselow Tape in overweight and obese pediatric patients in the Emergency Department. *Pediatric Emergency Care*. 2016;Oct 6. (DOI: 10.1097/PEC.0000000000000894; e-pub ahead of print).
21. Wells M, Goldstein L, Bentley A. A validation study of the PAWPER XL tape: accurate estimation of both total and ideal body weight in children up to 16 years of age. *Trauma Emerg Care* 2017;2:1-8.
22. Ramarajan N, Krishnamoorthi R, Strehlow M, Quinn J, Mahadevan SV. Internationalizing the Broselow tape: how reliable is weight estimation in Indian children. *Acad Emerg Med* 2008;15:431-6.
23. Asskaryar F, Shankar R. An Indian pediatric emergency weight estimation tool: prospective adjustment of the Broselow tape. *Int J Emerg Med* 2015;8:78.
24. Theron L, Adams A, Jansen K, Robinson E. Emergency weight estimation in Pacific Island and Maori children who are large-for-age. *Emerg Med Austr* 2005;17:238-43.
25. Sandell JM, Charman SC. Can age-based estimates of weight be safely used when resuscitating children? *Emerg Med J* 2009;26:43-7.
26. Shah AN, Frush K, Luo X, Wears RL. Effect of an 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.
27. Stevens AD, Hernandez C, Jones S, Moreira ME, Blumen JR, Hopkins E, et al. Color-coded prefilled medication syringes decrease time to delivery and dosing errors in simulated prehospital pediatric resuscitations: a randomized crossover trial. *Resuscitation* 2015;96:85-91.
28. Moreira ME, Hernandez C, Stevens AD, Jones S, Sande M, Blumen JR, et al. Color-coded prefilled medication syringes decrease time to delivery and dosing error in simulated Emergency Department pediatric resuscitations. *Ann Emerg Med* 2015.
29. Stewart D. Accuracy of the Broselow Tape for Estimating Paediatric Weight in Two Australian Emergency Departments. Thesis, University of Sydney, 2009. Available from: <https://ses.library.usyd.edu.au/bitstream/2123/6265/1/Declan%20Stewart%20WPP%20final.pdf> (Accessed 27 September 2017).
30. Altman D. *Practical Statistics for Medical Research*. London: Chapman and Hall, 1991.
31. Knight JC, Nazim M, Riggs D, Channel J, Mullet C, Vaughan R, et al. Is the Broselow tape a reliable indicator for use in all pediatric trauma patients?: A look at a rural trauma center. *Pediatr Emerg Care* 2011;27:479-82.
32. Akabarian S, Vahedparast H, Ravanipour M, Mirzaei K. Comparison of Broselow tape measurements versus mother estimations of pediatric weights. *Iranian South Med J* 2013;16:138-43.
33. Britnell S, Koziol-McLain J. Weight estimation in paediatric resuscitation: a hefty issue in New Zealand. *Emergency Med Australasia* 2015;27:251-6.
34. Abdel-Rahman SM, Jacobsen R, Watts JL, Doyle SL, O'Malley DM, Hefner TD, et al. Comparative performance of pediatric weight estimation techniques: a human factor errors analysis. *Pediatric Emergency Care*. 2017;33(8):548-552.
35. Rowe C, Koren T, Koren G. Errors by paediatric residents in calculating drug doses. *Arch Dis Childhood* 1998;79:56-8.
36. Potts M, Phelan K. Deficiencies on calculation and applied mathematics skills in pediatrics among primary care interns. *Arch Pediatr Adolesc Med* 1996;150:748-52.
37. Luscombe M, Owens B. Weight estimation in resuscitation: is the current formula still valid? *Arch Dis Child* 2007;92:412-5.
38. Cattermole GN, Leung PY, Graham CA, Rainer TH. Too tall for the tape: the weight of schoolchildren who do not fit the Broselow tape. *Emerg Med J* 2013;31:541-4.
39. Black K, Barnett P, Wolfe R, Young S. Are methods used to estimate weight in children accurate? *Emerg Med* 2002;14:160-5.
40. Chavez H, Peterson R, Lo K, Arel M. Weight estimation in an inner city Pediatric Emergency Department: the effect of obesity. *Am J Emerg Med* 2015;33:1364-7.
41. Garcia CM, Meltzer JA, Chan KN, Cunningham SJ. A validation study of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape-a new weight estimation tool. *J Pediatr* 2015;167:e1.
42. Truby H, Paxton SJ. The Children's Body Image Scale: reliability and use with international standards for body mass index. *Br J Clin Psychol* 2008;47 (Pt 1): 119-24.
43. Whitfield KC, Wozniak R, Pradinuk M, Karakochuk CD, Anabwani G, Daly Z, et al. Anthropometric measures are simple and accurate paediatric weight-prediction proxies in resource-poor settings with a high HIV prevalence. *Archives of Disease in Childhood* 2017;102:10-16.
44. Luten R, Wears RL, Broselow J, Croskerry P, Joseph MM, Frush K. Managing the unique size-related issues of pediatric resuscitation: reducing cognitive load with resuscitation aids. *Acad Emerg Med* 2002;9:840-7.
45. Frush KS. Medication errors in pediatric emergency care: developing a national standard? *Ann Emerg Med* 2007;50:369-70.