How and Why Paediatric Weight Estimation Systems Fail - A Body Composition Study

Mike Wells 1, Lara N. Goldstein 1

1. Emergency Medicine, University of the Witwatersrand, Johannesburg, ZAF

Corresponding author: Mike Wells, profmikewells@gmail.com

Abstract

Background

Weight estimation during medical emergencies in children is essential, but fraught with errors if the wrong techniques are used, which may result in critical drug dosing errors. Individualised weight estimation is required to allow for accurate dosing in underweight and obese children in particular. This study was designed to evaluate the associations between weight estimations from different systems and body composition in order to establish how and why they may perform well or poorly.

Methods

A convenience sample of 332 children aged from one month to 16 years had weight estimations using four age-based formulas: the Broselow™ Pediatric Emergency Tape (Armstrong Medical Industries, Inc., Lincolnshire, IL), the Mercy Method, and the Pediatric Advanced Weight Prediction in the Emergency Room, Extra-large/Extra-long Tape (PAWPER XL) Tape. They also had an assessment of body composition using dual x-ray absorptiometry (DXA). The weight estimates were compared against total body weight (TBW), calculated ideal body weight (IBW), and DXA-measured fat-free mass (FFM). Analyses of associations between age, length, weight estimation outcomes, and body composition were performed.

Results

Age-based formulas were very inaccurate because of the erratic relationship between age and body composition. The Broselow tape estimated IBW well in obese children because of the strong relationship between length and fat-free mass. It predicted TBW poorly in underweight and obese children, however, because of the poor relationship between length and fat mass. The Mercy Method’s performance was unrelated to body composition, but estimated TBW reasonably well and could not predict IBW or FFM. The PAWPER XL Tape’s performance was the most closely associated with body composition and, therefore, achieved an acceptable accuracy for estimations of TBW, IBW, and FFM.

Conclusions

Of the systems evaluated, the PAWPER XL Tape has the best association with body composition and the most accurate estimations of TBW, IBW, and FFM.

Introduction

The ultimate purpose of the weight estimation systems used during the management of medical emergencies in children is to enable the administration of accurate doses of potentially life-saving drugs [1]. If the weight estimation is inaccurate, or the methodology is prone to error when used during emergency situations, then the child is at risk of medical error [2]. An understanding of how and why these systems can fail would provide insight into which methodologies should be preferred and how they should be used within the context of the overall medical management of the critically ill or injured child. It would also provide understanding into how they could be improved. Malfunction could be as a result of inaccuracy of the estimation technique (producing a high incidence of critical errors), a failure to use the technique correctly (e.g., the Broselow™ Pediatric Emergency Tape used the wrong way round or calculation errors for formulas), a failure to function well in specific subgroups (e.g., underweight and overweight children), or a child falling outside of the weight estimation system’s parameters (e.g., children “too tall for the Broselow Pediatric Emergency Tape”) [3-4].

Some of the most commonly used weight estimation methods include formulas based on age, the Broselow Pediatric Emergency Tape, and newer dual length and habitus-based systems, such as the Mercy Method and

How to cite this article

Wells M, Goldstein L N (March 07, 2020) How and Why Paediatric Weight Estimation Systems Fail - A Body Composition Study. Cureus 12(3): e7198. DOI 10.7759/cureus.7198
the Pediatric Advanced Weight Prediction in the Emergency Room, Extra-large/Extra-long (PAWPER XL) Tape. Some of these methodologies, such as age-based formulas, have already been shown to be inaccurate in previous studies, but the underlying causes have not been well-studied [5-6]. Other techniques, such as the PAWPER XL Tape, have been shown to be accurate in some studies [7-9] but could still potentially be improved, especially as other studies have shown inconsistent accuracy in obese populations [10-11], primarily as a result of inaccurate habitus assessment.

The degree to which weight estimation systems can be considered to be inadequate depends on the standards by which they are judged. With the increasing prevalence of obesity in children, in both high-income as well as low- and middle-income countries, and a high prevalence of underweight children in low- and middle-income countries, weight estimation systems should ideally be able to provide sufficient information to allow correct dosing in children with both normal and extremes of body composition [4]. Although there is still some controversy about the correct dose scalar to use in obese children, the broad consensus is that, while total body weight (TBW) is still used for many drugs, ideal body weight (IBW) is required for the safe dosing of others [12]. This is important because there is some evidence that drug dosing errors might contribute to poorer outcomes in obese children suffering from cardiac arrest [13]. Therefore, while drug dose determination during emergencies may be difficult, a high standard of accuracy is required to ensure patient safety. For underweight children, an accurate estimation of TBW is essential, as IBW may be significantly higher than TBW [4]. Therefore, the ideal weight estimation system needs to be able to predict both TBW and IBW accurately, and an individualized plan should be employed for weight estimation in each child, dependent on their body composition or habitus.

The primary objectives of this study were to evaluate the vulnerabilities of selected age-based formulas, the Broselow Pediatric Emergency Tape, the Mercy Method, and the PAWPER XL Tape with regards to the prediction of an appropriate weight descriptor for drug dose calculations and to identify how variations in body composition could influence the accuracy of weight estimation.

**Materials And Methods**

This was a prospective, cross-sectional study conducted in an academic-aligned hospital in Johannesburg, South Africa, which treats approximately 6000 children per year. Ethics approval for this study was obtained from the Human Research Ethics Committee of the University of the Witwatersrand (approval number M120486).

A convenience sample of 332 children from one month to 16 years of age who presented to the Emergency Department (but did not require emergency treatment) were enrolled between October and December 2015. Exclusion criteria included failure to obtain consent and the inability to obtain critical measurements.

After basic demographic data were obtained, each child was changed into a hospital gown for the subsequent measurements. Anthropometric measurements of length, mid-arm circumference (MAC) and humerus length were obtained with the child in a supine position (to simulate emergency treatment conditions). Measured weight was obtained using a Tanita SC-240 Body Composition Analyser, following which whole-body dual x-ray absorptiometry (DXA) measurements of body composition were acquired using a Hologic Discovery A Densitometer (software version 12.6). All data were collected by one of the researchers (MW or LG).

TBW was estimated with the Broselow tape [14] and PAWPER XL tape [8] as well as the Mercy Method [15], the Advanced Paediatric Life Support (APLS) formulas, Erker formulas [16], the European Paediatric Life Support (EPLS) formula, and the Best Guess formulas (Table 1). A visual gestalt assessment of habitus was used to classify children for the Erker formulas into “thin,” “normal,” and “thick” categories [16]. Body mass index (BMI), BMI-for-age Z-scores, and an estimate of IBW (using the BMI50 method) were calculated using the Centers for Disease Control (CDC) growth charts [17].
## TABLE 1: A Description of the Weight Estimation Methods That Were Included in this Study

| Name                                      | Formula or method                                                                 | Restrictions                                      |
|-------------------------------------------|-----------------------------------------------------------------------------------|---------------------------------------------------|
| European Paediatric Life Support (EPLS)   | Wt = (2 × Z) + 8 or Wt = 2 × (Z + 4)                                              | Age restriction 1 to 10 years of age.              |
| formula (also known as the old APLS       | Wt = Z / 2 + 4                                                                    | For infants ≤12 months of age.                    |
| formula)                                   | Wt = (2 × Z) + 8 or Wt = 2 × (Z + 4)                                              | For children aged 1 to 5 years.                   |
|                                           | Wt = (3 × Z) + 7                                                                  | For children aged 6 to 12 years.                  |
| Advanced Paediatric Life Support (APLS)   | Wt = (2 × Z) + 8 or Wt = 2 × (Z + 4)                                              | For infants ≤12 months of age.                    |
| formula (also known as the new APLS       | Wt = (3 × Z) + 7                                                                  | For children aged 1 to 5 years.                   |
| formula)                                   |                                                                                  | For children aged 6 to 12 years.                  |
| Best Guess formula                         | Wt = (2 × Z) + 10 or Wt = 2 × (Z + 5)                                             | For children aged 1 to 5 years.                   |
|                                           | Wt = 4 × Z                                                                        | For children aged 6 to 14 years.                  |
|                                           | Wt = (2 × Z) + 6 or Wt = 2 × (Z + 3)                                              | For “thin” children aged 1 to 12 years.           |
| Erker formulas                             | Wt = (3 × Z) + 6 or Wt = 3 × (Z + 2)                                              | For “normal” children aged 1 to 12 years.         |
|                                           | Wt = (4 × Z) + 6                                                                  | For “thick” children aged 1 to 12 years.          |
| Broselow tape                              | An estimate of total body weight can be read directly off the tape laid next to a  | Length restriction 143 cm.                        |
|                                           | supine child.                                                                     |                                                   |
| PAWPER XL tape                             | An estimate of ideal body weight can be read directly off the tape laid next to a  | Length restriction 180 cm.                        |
|                                           | supine child. Total body weight can be estimated by adjusting the weight estimation up or down depending on a visual estimation of body habitus. The total body weight estimates can be read off the tape with no need for calculations. |
| Mercy method                               | Measurements of humeral length and mid-arm circumference are used to estimate weight. “Partial” weights are read off a table for each measurement and added for an estimation of total body weight.                  | For children aged 2 to 16 years.                  |

Fat-free mass (FFM), fat-free mass index (FFMI), and fat mass index (FMI) were derived from the DXA data for each child, using proprietary paediatric formulas installed by the manufacturer.
This data was evaluated using an analysis based on a modified Bland-Altman methodology. Each weight-estimation system was compared with TBW, IBW, and DXA-measured FFM using a percentage error analysis. Mean percentage error (MPE), 95% limits of agreement of percentage error, and the percentage of estimations falling within 10% (PW10) and 20% (PW20) of the weight descriptor were the three primary outcome measures used.

Subgroup analyses were performed in three weight categories (< 10 kg, 10 to 25 kg, and >25 kg) and three habitus categories based on BMI-for-age Z-scores (*thin* children Z-score ≤ 2.0, *fat* children Z-score ≥ 2.0, and normal-weight children in between). A subgroup analysis was also performed for those children falling outside the restrictions of the weight estimation systems. A PW10 of > 70% and a PW20 > 95% was considered to be an acceptable accuracy for a weight estimation method [2, 18].

The associations between age, length, and body composition were evaluated using a graphical representation and Pearson correlation analysis, with and without logarithmic transformation, as appropriate.

Data from each of the weight estimation systems were plotted on a Hattori chart, according to the percentage error category, to represent the associations between body composition and accuracy of weight prediction [19]. A description of the Hattori chart analysis is shown in Figure 1.

**FIGURE 1: An explanation of the Hattori chart analysis**

For each child in the study sample, the values of FM (kg) and FFM (kg) measured with DXA were used to calculate two new indices of body composition, normalised for height: FFMI (calculated as FFM/height²) and FMI (calculated as FM/height²). The values of FFMI and FMI for each child can then be plotted on the chart. The x-axis represents FFMI (kg/m²) and the y-axis FMI (kg/m²). The contributions of FFMI and FMI to BMI can be easily identified. Isolines of BMI and % body fat are represented by diagonal lines across the chart (in blue and red, respectively). In the absence of adequate reference data for children with respect to FFMI and FMI, the first and third quartiles for FFMI and FMI from the current study population were plotted on the chart. These are represented by the vertical and horizontal black lines, respectively. Four kinds of information can be read off the chart at the same time: FMI, FFMI, BMI, and % body fat, in addition to identifying any individual child’s FFMI and FMI relative to the study population.

BF: body fat; BMI: body mass index; DXA: dual x-ray absorptiometry; FFMI: fat-free mass index; FM: fat mass; FMI: fat mass index; IQR: interquartile range
In order to evaluate the accuracy of the assignment of habitus score (HS) for the PAWPER XL, an "ideal" HS was determined for each child (the HS which would result in the best weight estimate at the child’s length), which was compared to the original HS. For comparisons with IBW and FFM, the PAWPER XL HS3 and HS1 weights, respectively, were used.

**Results**

**Characteristics of study participants**

A total of 332 children were included in the study. The basic demographic and body composition data is shown in Table 2. There were a substantial number of underweight children (15.3%) and overweight or obese children (22.3% and 10.2%, respectively). This allowed for the weight estimation systems to be tested over a spectrum of body habitus variations. Subgroup data for thin, normal, and fat children were based on categories derived using the BMI system described by Cole et al. [20]. Since body composition reference data have not been well established in children, and especially in younger children, most analyses in this study made use of pragmatic limits that might affect drug dosing decisions: children were considered to be significantly "fat" when their TBW was > 120% of IBW (which roughly corresponds to a Z-score of +2) as this would require the use of IBW as a drug dosing descriptor for some drugs. Likewise, children were considered significantly "thin" when their IBW was > 120% TBW (which approximates a Z-score of -2) as the use of IBW would result in a critical overdosing error. There was a clinically important number of children whose TBW and IBW differed by more than 20%.

![Table 2: Description of the Study Population: Demographic Information with Body Composition Data](image-url)
Main results

The performances of each of the weight-estimation systems against TBW, IBW, and DXA-measured FFM are shown in Table 3. With regards to estimating TBW, the overall accuracy of the regular age-based formulas was very poor (PW10 ranging from 29.7 to 43.1%) with even the habitus-modified Erker formula achieving a PW10 of only 50.0%. The Broselow tape and the Mercy method achieved an intermediate degree of accuracy (PW10s of 52.1 and 63.9%, respectively) and the PAWPER XL tape achieved a high degree of accuracy (PW10 83.4%). The age-formulas and the Mercy method did not estimate IBW accurately in obese children (PW10IBW 24.1% to 39.7% and 12.1%, respectively). The Broselow tape and the PAWPER XL tape (using the HS weight to predict IBW) both predicted IBW well (PW10IBW 72.4 and 87.9%, respectively), with the PAWPER XL tape accurately identifying all the obese children (those scored as an HS > 5). None of the other systems had a mechanism to identify obese children. When it came to predictions of FFM, only the PAWPER XL tape (using the HS1 weight to predict FFM) and the EPLS formula achieved a PW10IBW of greater than 50%, with PW10s of 82.4% and 58.6%, respectively.

|                      | Total body weight (TBW) | Ideal body weight (IBW) | Measured FFM (SBW) |
|----------------------|-------------------------|-------------------------|---------------------|
|                      | PW10       | PW30       | MPE    | LLOA    | ULLOA | PW10        | PW30      | MPE         | LLOA     | ULLOA     |
| All                  | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| < 10 kg              | 101.1      | 97.6       | -3.5   | -20.3   | 18.9   | 6.6        | 21.9      | 31.1       | -5.1     | 66.6      |
| 10-25 kg             | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| > 25 kg              | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| Broselow tape        | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| Normal               | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| Fat                  | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| > 145 cm             | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| HS > 5               | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| Mercy method         | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| Normal               | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| Fat                  | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| > 145 cm             | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| HS > 5               | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| PAWPER XL tape TBW   | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| Normal               | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| Fat                  | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| > 145 cm             | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| HS > 5               | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| PAWPER XL tape IBW   | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
| Obese                | 102.0      | 98.2       | -4.5   | -21.8   | 19.3   | 6.6        | 22.0      | 31.9       | -4.9     | 62.7      |
For the purposes of this evaluation, children were defined as “thin” if TBW was less than 90% of IBW, “fat” if TBW was greater than 120% of IBW.

The data is presented for the whole sample as well as by subgroups of weight, habitus, and two special categories (length > 145 cm and HS > 5).

### Accuracy Outcomes of the Weight Estimation Systems Evaluated in This Study

| System                | All          | < 10 kg      | 10-25 kg     | > 25 kg      | < 10 kg      | 10-25 kg     | > 25 kg      |
|-----------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| PAMPER XL, taped LBW  | 232 (42.1%)  | 11 (27.3%)   | 168 (45.8%)  | 153 (35.9%)  | 11 (27.3%)   | 168 (45.8%)  | 153 (35.9%)  |
| EPLS formula (European Paediatric Life Support formula) | 28 (4.6%)    | 0 (0.0%)     | 28 (4.6%)    | 0 (0.0%)     | 28 (4.6%)    | 0 (0.0%)     | 28 (4.6%)    |
| APLS formula (new Advanced Paediatric Life Support formula) | 15 (23.1%)   | 5 (0.0%)     | 10 (25.5%)   | 13 (20.0%)   | 5 (0.0%)     | 10 (25.5%)   | 13 (20.0%)   |
| PAWPER XL tape LBW   | 34 (6.1%)    | 4 (2.0%)     | 26 (6.1%)    | 31 (6.2%)    | 4 (2.0%)     | 26 (6.1%)    | 31 (6.2%)    |
| HS > 5                | 15 (23.1%)   | 5 (0.0%)     | 10 (25.5%)   | 13 (20.0%)   | 5 (0.0%)     | 10 (25.5%)   | 13 (20.0%)   |
| > 145 cm              | 15 (23.1%)   | 5 (0.0%)     | 10 (25.5%)   | 13 (20.0%)   | 5 (0.0%)     | 10 (25.5%)   | 13 (20.0%)   |

### TABLE 3: Accuracy Outcomes of the Weight Estimation Systems Evaluated in This Study

The data is presented for the whole sample as well as by subgroups of weight, habitus, and two special categories (length > 145 cm and HS > 5). For the purposes of this evaluation, children were defined as “thin” if TBW was less than 90% of IBW, “fat” if TBW was greater than 120% of IBW.
The PAWPER XL tape, the Mercy method, and the APLS formula were able to provide valid weight estimations for all children in the study sample, with the Broselow tape, the EPLS, Erker, and Luscombe formulas outside of their restrictions in 14.5%, 10.5%, 10.5%, and 22.9%, respectively. For the subgroup of children "too tall" for the Broselow tape (those with a length > 145 cm), only the Mercy method and the PAWPER XL tape maintained acceptable accuracy.

The relationships between age, length, and body composition components (including FFM and FM components) are shown in Figure 2 (Panels A-D). The relationship between age and weight (Panel A) was weaker with TBW than with IBW ($r^2 = 0.74$ and $r^2 = 0.91$ for TBW and IBW, respectively, $p < 0.001$). This difference was greatest in older children. Once the effect of height was removed from the age (Panel B) by expressing FFM and FM as indices (FFM/height$^2$ and FM/height$^2$), the very poor association between age and FFM and FM was exposed ($r^2 = 0.26$ and $r^2 = 0.03$, respectively, $p < 0.05$ and $p = 0.8$, respectively). Length and weight were strongly correlated (Panel C), although the relationship between TBW and length was weaker than between IBW and length ($r^2 = 0.77$ and 0.93 for TBW and IBW, respectively, $p < 0.001$). This association was reflected in the association between length and FFM and FM (Panel D). Length was strongly correlated with FFM ($r^2 = 0.86$, $p < 0.001$) but far less well with FM ($r^2 = 0.48$, $p < 0.001$). This relationship held true when evaluating length against logarithmic transformations of FFM and FM ($r^2 = 0.96$ and $r^2 = 0.66$, respectively, $p < 0.001$). In summary, the relationship between age and TBW and IBW was much weaker than that between length and weight descriptors. While the length and IBW and the length and FFM were strongly correlated, the relationship with TBW and FM was significantly weaker.
Figure 3 (Panels A to D) illustrates the associations between body composition and the accuracy of the age-based weight estimation systems in the Hattori chart format. The EPLS formula (Panel A) was the most accurate in children with a low BMI, with no discernible difference in discriminating between FMI or FFMI. In contrast to this, the APLS formula (Panel B) and the Best Guess formula (Panel C) showed a substantially greater overestimation of weight but could not differentiate between children with low and high BMI. The Erker method (Panel D) had a far greater accuracy at all values of BMI than the other formulas. It also, however, produced some large overestimations of weight, even in children of normal weight. In summary, the age-based formulas were all inaccurate but with different biases. The EPLS formula underestimated weight in higher-BMI children, while the APLS and Best Guess formulas overestimated the weight of low-BMI children. The Erker formula had less relationship with body composition but still failed to predict weight accurately.

**FIGURE 3:** Hattori charts of the study population showing outcomes of total body weight estimation by the age-based formulas

Panel A: outcomes of total body weight estimation by the European Paediatric Life Support (EPLS) age-based formula; Panel B: outcomes of total body weight estimation by the new Advanced Paediatric Life Support formula (APLS) age-based formula; Panel C: outcomes of total body weight estimation by the Best Guess age-based formula; Panel D: outcomes of total body weight estimation by the Erker age-based formula.

Square markers represent an overestimation of weight and round markers represent an underestimation of weight. Markers with a green fill indicate a weight estimation accuracy of within 10% of actual weight; orange markers an accuracy of between 10% and 20% of actual weight; and red markers an error of greater than 20%. The medians for each error category are shown in black.

BF: body fat; BMI: body mass index; FFMI: fat-free mass index; FMI: fat mass index

Figure 4 (Panels A to C) shows the Hattori charts of the length and dual length- and habitus-based methods. The Broselow tape’s performance (Panel A) showed good accuracy in children in the central ranges of FFMI and FMI, but poor results at the extremes of habitus. Both the PAWPER XL tape (Panel B) and the Mercy tape (Panel C) demonstrated better accuracy in the central ranges of BMI and FMI compared to the Broselow tape.
method (Panel C) revealed a substantially better performance than the other methods at all ranges of FFMI and FMI. While the PAWPER XL tape still showed some inaccuracies at the far extremes, especially of FMI, the Mercy method’s performance was virtually independent of habitus, even at the extremes. In summary, these methods showed fewer critical errors in weight estimation than age-formulas. The Broselow tape showed errors in all children with higher or lower than average BMI, while the PAWPER XL tape showed the same errors only at the extremes of habitus. The performance of the Mercy method showed almost no relationship to body habitus or composition.

FIGURE 4: A Hattori chart of the study population showing outcomes of total body weight estimation by the length and length- and habitus-based methods

Panel A: outcomes of total body weight estimation by the Broselow tape; Panel B: outcomes of total body weight estimation by the PAWPER XL tape; Panel C: outcomes of total body weight estimation by the Mercy method

Square markers represent an overestimation of weight and round markers represent an underestimation of weight. Markers with a green fill indicate a weight estimation accuracy of within 10% of actual weight; orange markers an accuracy of between 10% and 20% of actual weight; and red markers an error of greater than 20%. The medians for each error category are shown in black.

BF: body fat; BMI: body mass index; FFMI: fat-free mass index; FMI: fat mass index; PAWPER XL tape: Paediatric Advanced Weight Prediction in the Emergency Room extra large/extra long tape

Figure 5 (Panels A and B) contains Hattori charts showing the distribution of habitus scores assigned to the study population. The data showed a good correlation between FMI and HS ($r^2 = 0.70$) and a lesser correlation with FFMI ($r^2 = 0.43$). It also shows a comparison between the actual HS assigned in this study and the "ideal" post hoc HS. In this analysis 51.2% of scores remained unchanged, 41.6% differed by one HS point and the remaining 7.2% differed by two points. In summary, habitus scores accurately represented differences in body composition, especially changes in FM. Ideal and actual habitus score assignments were similar except at extremes of obesity.
FIGURE 5: A Hattori chart of the study population showing outcomes of habitus score assignment for the PAWPER XL tape

Panel A: the actual habitus score assignments. Orange markers represent HS1 and HS2; green markers represent HS3; yellow markers represent HS4 and red markers represent HS5 and above. The coloured ellipses approximate the 95% limits of agreement for the distribution of each habitus score group; Panel B: a Hattori chart of the FFMI and FMI medians of assigned habitus scores for the PAWPER XL tape. The chart shows the median FFMI and FMI of the actual (blue) and ideal (green) habitus scores (from HS1 to HS7, left to right).

BF: body fat; BMI: body mass index; FFMI: fat-free mass index; FMI: fat mass index; HS: habitus score; PAWPER XL tape: Paediatric Advanced Weight Prediction in the Emergency Room extra large/extra long tape

Discussion
Age and length as predictors of TBW, IBW, FFM, and FM

The relationship between age and weight and length and weight is fundamental to the ability of these variables to predict weight. Length was clearly more closely associated than age with TBW, as has been shown previously [21], as well as with IBW and FFM. Once length was removed as a confounding variable, the effect size of the correlation between age and weight was seen to be small. In terms of length, the association with FFM was much closer than with FM, which further explains why one dimensional length-based systems (such as the Broselow tape) were able to predict IBW well but did not predict TBW well because of the potential large variations in FM at any given length [22]. Although length was demonstrably superior to age as a weight-estimation variable, some method of assessing habitus is clearly also required to account for differences in FM between children.

A note of caution: although FFM, lean body weight (LBW), and IBW are frequently used interchangeably as an appropriate scalar for drug dosing in obese children, they are not identical. IBW has no true biological validity but has been shown to be similar to FFM in older children [23]. The relationship between LBW and FFM is also unclear in children, but LBW is probably 5% to 10% higher than FFM and is the true "pharmacological scalar" that is desired for calculating doses for hydrophilic medications in obese children [24]. How this should best be translated into clinical practice is, as yet, undetermined.

Interpreting the Hattori chart analysis of weight estimation systems

For the unmodified age-based formulas, the EPLS formula was reasonably accurate only in children with an FFMI and FMI at the lower end ranges; the opposite was true of the Best Guess formula, with the APLS formula falling somewhat between the two. Differences between the age formulas were, essentially, whether the major errors were greater in lower BMI children (APLS and Best Guess formulas) or in higher BMI children (EPLS formula). The degree of weight estimation error for the formulas was only poorly associated with body composition, supporting the findings that age is much less predictive of body weight than length. The habitus-modified Erker formula, although more accurate than the other formulas, showed the most random association between body composition and accuracy. This strongly suggests that, despite a good theoretical basis for this system, it will not be able to achieve adequate accuracy.

The Broselow tape: the most accurate weight estimations were in the central zone (the "normal weight" child), with critical inaccuracies at increasing FFMI and FMI and decreasing FFMI. This was typical of what might be expected - a good relationship between length and weight but variations in FFM/FM not accounted for [22]. Most of the inaccurate estimations were underestimations (from higher than average BMI, and to a lesser extent, FFMI). When considering what the body composition analysis revealed in terms of how the various methods failed, it was clear that the age formulas and Broselow tape were unable to produce accurate weight estimations outside of their narrow area of calibration. Unless these methods are restricted to patient populations that fall within these narrow limitations, they should not be used if better methods are available.

The Mercy method: this method’s performance showed almost no association with body composition. Almost all error categories had FFMI/FMI medians within the central region, with the exceptions of the few children with > 10% weight overestimation who were in the high FFMI/FMI sector. This pattern indicated that some other factor, unrelated to habitus, was the root cause of weight estimation error, such as measurement errors, regional aberrations in body composition, or unknown factors. While the Mercy method was able to accurately predict weight in children of all different body compositions, its failures bore no relation to extremes of body composition. This was, therefore, not a "calibration" error, which would make it difficult to identify specific vulnerabilities and devise methods to improve accuracy.

The PAWPER XL tape: this method showed poor accuracy only at extremes of FFMI and FMI. This was in contrast to the Broselow tape, which showed poor estimation at smaller deviations from the medians. This indicated that the PAWPER XL tape was better "calibrated" as it accounted for a greater degree of variation in body composition. The pattern of inaccuracies further indicated that the tape could potentially be improved with better weight estimation at extremes of body habitus: "fatness" (high FMI) was underestimated in the higher habitus scores, but "slimness" (low FFMI) was under-recognised in the lower habitus scores.

Putting it all together

Age-based Formulas Failed Badly Because of Large Variability of Weight-for-age (FFM and FM)

In this study, the strength of the association between age and TBW or IBW was intermediate at best. This was because the association of age with body composition (FFMI and FMI) was very weak, which was evident once length was removed as a confounding variable. This strongly suggests that it is unlikely that age can ever be used to accurately estimate weight and will always be inferior to length-based systems [25]. The association of age with both FFMI and FMI was weak enough to suggest that neither TBW nor IBW would be able to be accurately predicted. Although the formulas were most accurate in the 10 to 25 kg category (as has previously been found), this accuracy never achieved acceptable levels (PW10 > 70% and PW20 > 95%) as the
variability of FFM and FMI remained consistent across the age range of the population [26]. No age formula has ever been shown to perform satisfactorily well in any previous study (best performances: PW10 45% to 55%) [2]. The new Erker formula, which is habitus-modified, failed to deliver on the potential showed in its theoretical development [16]. This error was predominantly one of modest overestimation of weight in children with normal habitus and modest underestimation of weight in fat children. There were, however, relatively fewer critical errors (estimation error > 20%) than with the other formulas. The biggest weakness with this system was that, at any age, children could have the same habitus but very different lengths (and, therefore, weights) that cannot be accounted for by this system. It is unlikely that further calibration will improve this system significantly.

The Broselow Tape Fails Because of Large Variability of FM for Length

Numerous studies from across the world have confirmed that the tape overestimates weight in underweight populations (often to a dangerous degree) and underestimates weight in populations with a high prevalence of obesity [2, 26]. Although the relationship between length and weight was far stronger than that between age and weight, the relationship between length and FM was far more inconsistent and increasingly variable with increasing length. This variability was sufficient to account for the relatively poor performance of the tape. It also explains why the tape predicted TBW well in children of "normal" habitus, but poorly in all others. This is why researchers have successfully proven the possibility of using alternative methods of habitus assessment to improve the accuracy of the Broselow tape [27, 28].

The present study found that the relationship between length and FFM was strong, which explained the excellent association between length and IBW for the Broselow tape. Importantly, however, this is only relevant for obese children (with respect to drug dosing for hydrophilic drugs), but the Broselow tape as yet has no validated mechanism for identifying obese children. This needs to be considered if the Broselow tape is used in overweight or obese children.

The Mercy Method Fails Because of Non-habitus-related Factors

Generally, the Mercy method predicted TBW well across the age and habitus spectrum but was weakest in infants and obese children. Critical errors were very uncommon. It predicted IBW poorly, especially in obese children in whom IBW would be required, but this was not surprising, as the method was not designed to predict IBW. The system also had no mechanism to identify obese children for whom an alternate weight-descriptor (such as IBW) might be required. The inaccurate estimations incurred by this method were not specifically attributable to habitus, although the poorest accuracy was in obese children (with a mixture of under- and overestimation of weight, however). Part of the explanation might be that, in this study, children were measured in the supine position (simulating how it would be used in medical emergencies), unlike previous studies, which might have led to less accurate measurements [29]. Other studies, not by the developers of the system, have also shown a substantially poorer accuracy than in the original studies [2, 7, 9]. Nonetheless, the Mercy method remains one of the most accurate systems available today, although not ideally suited for use in the Emergency Department.

The PAWPER XL Tape Fails Because of Incorrect Assignment of Habitus Scores

Generally, the PAWPER XL tape predicted TBW, IBW, and FFM extremely well but was weakest in infants (overestimating weight) and morbidly obese children (underestimating weight). The rate of critical estimation errors for all weight descriptors was the lowest amongst all the weight estimation systems. The errors of estimation appeared to be directly related to body habitus, which suggested that an improvement in the assessment of habitus might further improve the tape's "calibration" and, therefore, accuracy. Studies in obese populations in the United States have also demonstrated this vulnerability to inaccurate HS assignment [11]. The "ideal" HS assignments showed that the habitus estimations were frequently more moderate than the ideal scores and that it was theoretically possible to achieve virtually perfect weight estimation with ideal habitus assessment. However, it is unclear whether it is possible in practice to differentiate between the relatively minor differences in external indicators of body composition that might be required to achieve this accuracy. Validated reference figural images will likely be required to enable the habitus assessment to be standardised and generalised across different populations. This technique has been validated in preliminary studies for the PAWPER XL tape, but further research is needed [30].

The ability of the PAWPER XL tape to accurately estimate FFM and IBW, in addition to being able to identify the obese children for whom this would be required for dosing calculations, was an interesting finding which might prove useful in the future. This requires further research.

Conclusions

The association between predictive variables (age and length) and TBW, IBW, FM, and FFM showed the underlying biological limitations of using age or length alone to predict weight. None of the age-based formulas achieved a satisfactory degree of accuracy, including the age-based habitus-modified formulas of Erker. Differences between the age formulas were essentially whether the major errors were greater in
lower-BMI children (APLS and Best Guess formulas) or in higher-BMI children (EPLS formula). The Broselow tape performed better than the age-formulas but it did not achieve satisfactory accuracy in estimating TBW. It was accurate in predicting IBW, however, and could be used for this purpose for drug dose calculations as long as TBW was known or estimated using another technique. The Mercy method (a dual length- and habitus-based method) demonstrated a higher accuracy in estimating TBW than the univariate methods, with fewer critical errors. Its performance was completely independent of body composition. This suggested that other patient factors or user errors might be critical determinants of its functioning accurately. It was unable to estimate IBW. The performance of the dual length- and habitus-based PAWPER XL tape was the best of all the systems. It was also the only system that had a specific mechanism to produce estimations of TBW (the HS3 weight) and FFM (the HS1 weight). The overall accuracy of estimation of TBW, IBW, and FFM were very good, even for children outside of the restrictions of other methods. Its weakest performance was in children with extreme habitus types, which was probably mostly due to errors or aberrations in the assessment of body habitus. This needs to be addressed in future research.

**Additional Information**

**Disclosures**

**Human subjects:** Consent was obtained by all participants in this study. Human Research Ethics Committee, University of the Witwatersrand issued approval M120486. Approved unconditionally. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

**References**

1. Young KD, Korotzer NC: Weight estimation methods in children: a systematic review. Ann Emerg Med. 2016, 68:441-51. 10.1016/j.annemergmed.2016.02.043
2. Wells M, Goldstein L, Bentley A: The accuracy of emergency weight estimation systems in children - a systematic review and meta-analysis. Int J Emerg Med. 2017, 10:29. 10.1186/s12245-017-0156-5
3. Benjamin L, Frush K, Shaw K, et al.: Pediatric medication safety in the emergency department. Pediatrics. 2018, 141:e20174066. 10.1542/peds.2017-4066
4. Wells M, Goldstein LN: Optimising emergency weight estimation in underweight and obese children: the accuracy of estimations of total body weight and ideal body weight. Eur J Emerg Med. 2019, 26:301-307. 10.1097/MEJ.0000000000000561
5. Marlow R, Lo D, Walton L: Accurate paediatric weight estimation by age: mission impossible?. Arch Dis Child. 2011, 96:A1-2. 10.1136/adc.2011.212563.3
6. Marlow RD, Wood DLB, Lyttle MD: Comparing the usability of paediatric weight estimation methods: a simulation study. Arch Dis Child. 2019, 104:121-25. 10.1136/archdischild-2018-314873
7. Georgoulas V, Wells M: The PAWPER tape and the Mercy Method outperform other methods of weight estimation in children in South Africa. S Afr Med J. 2016, 106:935-39. 10.7196/SAMJ.2016.v106i9.10572
8. 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. 10.15761/TEC.1000141
9. Shrestha K, Subedi P, Pandey O, Shakya L, Chhetri K, House DR: Estimating the weight of children in Nepal by Broselow, PAWPER XL and Mercy method. World J Emerg Med. 2018, 9:276-81. 10.5847/wjem.j.1920-8642.2018.04.007
10. Chavez H, Peterson R, Lo K, Arel M: Weight estimation in an inner-city pediatric ED: the effect of obesity. Am J Emerg Med. 2015, 33:1364-67. 10.1016/j.ajem.2015.07.040
11. 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, 167:173-7. e1. 10.1016/j.jpeds.2015.04.006
12. WHO Model Formulary for Children 2010. (2010). Accessed: February 6, 2020: http://apps.who.int/iris/handle/10665/44309.
13. Srinivasan V, Nadkarni VM, Helfaer MA, Carey SM, Berg RA; American Heart Association National Registry of Cardiopulmonary Resuscitation Investigators: Childhood obesity and survival after in-hospital pediatric cardiopulmonary resuscitation. Pediatrics. 2010, 125:481-88. 10.1542/peds.2009-1524
14. 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 Emerg Med. 1988, 17:576-81. 10.1016/0196-0644(88)90396-2
15. Abdel-Rahman SM, Ridge AL: An improved pediatric weight estimation strategy. Open Med Devices J. 2012, 4:87-97. 10.2174/18751814014040410087
16. Erker CG, Santamaria M, Moellmann M: Size does matter—age-related weight estimation in ‘tall n’ thin’ and ‘tiny n’ thick’ children and a new habitus-adapted alternative to the EPLS-formula. Resuscitation. 2014, 85:1174-78. 10.1016/j.resuscitation.2014.04.052
17. Phillips S, Edbeck A, Kirby M, Goday P: Ideal body weight in children. Nutr Clin Pract. 2007, 22:240-45. 10.1177/0884533607298022
18. Stewart D: Accuracy of the Broselow tape for estimating paediatric weight in two Australian Emergency Departments: University of Sydney (Thesis). (2009). Accessed: February 7, 2020.
19. Hattori K, Tatsunami N, Tanaka S: Assessment of body composition by using a new chart method. Am J Hum Biol. 1997, 9:573-78. 10.1002/(SICI)1520-6500(199709)9:5<573::AID-AJHB5>3.0.CO;2-V
20. Cole T, Bellizzi M, Flegal K, Dietz WH: Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ. 2000, 320:1240-43. 10.1136/bmj.320.7244.1240
21. Goldman JL, Ojoo A, Abdel-Rahman SM: Challenges in pediatric oral dosing. Optimizing Treatment for Children in the Developing World, 2015th Edition. MacLeod S, Hill S, Koren G, Rane A (ed): Springer International Publishing, Cham; 2015. 33:44. 10.1007/978-3-319-15750-4_4
22. Wells M, Goldstein L, Bentley A, Baanett S, Monteith I: The accuracy of the Broselow tape as a weight estimation tool and a drug-dosing guide - a systematic review and meta-analysis. Resuscitation. 2017, 121:9-33. 10.1016/j.resuscitation.2017.09.026
23. Ross EL, Heizer J, Mixon MA, Jorgensen J, Valdez CA, Czaja AS, Reiter PD: Development of recommendations for dosing of commonly prescribed medications in critically ill obese children. Am J Health Syst Pharm. 2015, 72:542-56. 10.2146/ajhp140280
24. Wells M, Goldstein L, Bentley A: Development and validation of a method to estimate bodyweight in critically ill children using length and mid-arm circumference measurements - the PAWPER XL-MAC system. S Afr Med J. 2017, 107:1015-21. 10.7196/SAMJ.2017.v107i11.12505
25. Wells M, Goldstein L, Bentley A: It is time to abandon age-based emergency weight estimation in children! A failed validation of 20 different age-based formulas. Resuscitation. 2017, 116:73-83. 10.1016/j.resuscitation.2017.05.018
26. Wells M, Goldstein LN, Bentley A: A systematic review and meta-analysis of the accuracy of weight estimation systems used in paediatric emergency care in developing countries. Afr J Emerg Med. 2017, 7:536-54. 10.1016/j.ajem.2017.06.001
27. 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. Pediatr Emerg Care. 2017, 33:675-78. 10.1097/PEC.0000000000001894
28. Wells M, Goldstein L, Bentley A: Accuracy of weight estimation by the Broselow tape is substantially improved by including a visual assessment of body habitus. Pediatr Res. 2018, 83:85-92. 10.1038/pr.2017.222
29. Wells M, Goldstein L, Bentley A: The accuracy of paediatric weight estimation during simulated emergencies: the effects of patient position, patient cooperation and human errors. Afr J Emerg Med. 2018, 8:45-50. 10.1016/j.ajemer.2017.12.003
30. Wells M, Goldstein L, Bentley A: The use of body habitus reference images improves the ability of novices to accurately estimate children’s weight using the PAWPER XL tape system. J Emerg Med. 2018, 54:165-75. 10.1016/j.jemermed.2017.10.009