Accuracy of Whole-Body Plethysmography Requires Biological Calibration

Porntiva Poorisrisak, Carsten Vrang, Jorn Molgaard Henriksen, Bent Klug, Birgitte Hanel and Hans Bisgaard

*Chest* 2009;135;1476-1480
DOI 10.1378/chest.08-1555

The online version of this article, along with updated information and services can be found online on the World Wide Web at:
http://chestjournal.chestpubs.org/content/135/6/1476.full.html

*Chest* is the official journal of the American College of Chest Physicians. It has been published monthly since 1935. Copyright 2009 by the American College of Chest Physicians, 3300 Dundee Road, Northbrook, IL 60062. All rights reserved. No part of this article or PDF may be reproduced or distributed without the prior written permission of the copyright holder.

(http://chestjournal.chestpubs.org/site/misc/reprints.xhtml)
ISSN:0012-3692
Accuracy of Whole-Body Plethysmography Requires Biological Calibration*

Porntiva Poorisrisak, MD; Carsten Vrang, MD; Jorn Molgaard Henriksen, DMSci; Bent Klug, DMSci; Birgitte Hanel, DMSci; and Hans Bisgaard, DMSci

Background: Specific airway resistance (sRaw) measured by whole-body plethysmography in young children is increasingly used in research and clinical practice. The method is precise and feasible. However, there is no available method for calibration of the resistance measure, which raises concern of accuracy. Our aim was to determine the agreement of sRaw measurements in six centers and expand normative sRaw values for nonasthmatic children including these centers.

Method: Identical hardware with different software versions was used at the six centers. Measurements followed a standard operating procedure: (1) seven healthy young children were brought to each of the six centers for sRaw measurements; and (2) 105 healthy preschool children (52 boys; mean age, 5.1 years; interquartile range, 4.3 to 6.0) were recruited locally for sRaw measurements.

Results: (1) The sRaw of the seven-children study group was significantly lower at two centers compared with the other four centers, and one center had significantly higher sRaw than all the other centers (p < 0.05). Error in the factory settings of the software was subsequently discovered in one of the deviating centers. (2) Normative data (105 preschool children) were generated and were without significant difference between centers and independent of height, weight, age, and gender. We subsequently pooled these normative data (105 children) with our previous data from 121 healthy young children (overall mean sRaw, 1.27; SD, 0.25).

Conclusion: Control using biological standards revealed errors in the factory setting and highlights the need for developing methods for verification of resistance measures to assure accuracy. Normative data were subsequently generated. Importantly, other centers using such normative data should first consider proper calibration before applying reference values.

(CHEST 2009; 135:1476–1480)

Key words: lung function tests; multicenter study; quality control; specific airway resistance; whole-body plethysmography

Abbreviations: BTPS = body temperature and pressure, saturated; P = pressure; sRaw = specific airway resistance; V = airflow

We have introduced and documented whole-body plethysmography for measurement of specific airway resistance (sRaw) in young children during the recent decade.1–5 sRaw assesses the airway resistance from measurements of the pressure changes driving the airflow (V) during tidal breathing. These measurements require no active cooperation and are therefore feasible in children from 2 years of age. sRaw is now increasingly used in both research and clinical practice.

We previously documented the precision (interobserver variability)4 of sRaw measurements, but the accuracy of the method has not been reported. Using reference values generated by other centers is particularly vulnerable to the accuracy of the methods used.2 Attempts have been made to develop a mechanical infant lung model analog for quality control of a whole-body infant plethysmograph,6,7 but it turned out to be difficult because of the small pressure and flow changes and is not readily available. Therefore, it is of concern that the accuracy of sRaw measurements cannot be verified. Flow and box leak are checked routinely, but the composite resistance measure is generated by algorithms buried in the software with settings often inaccessible to the end user. Thus, errors in
software or mechanics could go unnoticed with a potential impact on clinical evaluation.

The aim was to study center agreement by comparing sRaw measurements in young children among six centers in Denmark currently using the same equipment. First, we used seven healthy young children as a biological standard and brought them to each center for sRaw measurements. Second, we recruited healthy young children at five of the centers to compare and expand normative data.

**Materials and Methods**

The study was approved by the local ethics committee as a quality assurance project and approved by the Danish Data Protection Agency. Parents gave written informed consent.

**Design**

Seven healthy young children were recruited for the measurements at six participating centers. The children were measured at each center by a center-specific observer as well as an observer visiting each center (Dr. Poorisrisak). The two observers were blinded to each other’s measurements. Measurements in the individual children were finished within a period of 3 months. The order of center visits was randomized.

Healthy preschool children were recruited by random selection through the Central Person Registry from the local catchment area of the five centers. Children included were born at term, with no history of asthma-related symptoms, other chronic lung symptoms, or use of asthma treatment. If the child had a lower respiratory tract infection within the week before the appointment, the measurement was rescheduled. The children attended their local center where duplicate measurements were done by a local observer.

**Principle of Measurement**

Measurements were conducted in a constant volume whole-body plethysmograph (Master Screen Body; Erich Jaeger GmbH; Würzburg, Germany). A transducer measured pressure changes in this sealed box, and a pneumotachograph simultaneously measured the flow swing at the mouth.

$s_{Raw}$ was calculated as the ratio between the pressure ($P$) generated by thoracic and abdominal movements during tidal breathing and the resulting $V$.

$$s_{Raw} = \frac{\Delta P}{\Delta V}$$

where $\Delta P$ is the change in $P$ and $\Delta V$ is the change in $V$, in comparison with resistance in ohms.

Flow and volume measurements were corrected to body temperature and pressure, saturated (BTPS) with water vapor conditions, as follows:

$$s_{Raw} = (\Delta P/\Delta V) \times (P_{amb} - P_{H2O})$$

where $P_{amb}$ is ambient pressure and $P_{H2O}$ is pressure of water vapor at body temperature. The equipment was calibrated daily for ambient conditions (room temperature, atmospheric pressure, and humidity), box calibration (leak test result should be between 4 and 7 s and test for internal pressure, which should result in a correction factor of $<3\%$), and volume calibration (piston was pulled regularly 10 times with a 3-L piston and automatically accepted or rejected by the software). All the centers used identical hardware, but software versions differed between centers (JLAB, versions 4.51, 4.53 with different subversions, 4.65, and 4.67; available at http://www.viasyshealthcare.com).

**Procedure of Measurement**

The same procedure was followed by all observers. The children were seated alone in the box with the door closed. The child’s breathing aimed for a frequency of 30 to 45 breaths/min.°

The children used a face mask with a large cushion, which ensured a good seal and stabilized the cheeks and chin. A built-in flexible tube ensured that the mouth remained open to avoid nasal breathing.

“Loops” on the screen showed the relation between pressure (or volume) [x-axis] and flow (y-axis) [i.e., the pressure driving the air flow in and out of the lungs]. $s_{Raw}$ was estimated from the inclination of these loops using the line between points of maximum pressure (or $s_{RawTOT}$).

Technically acceptable loops were chosen as those that were “closed” in the middle. “Open” loops normally indicated insufficient BTPS correction. The loops assumed a straight line with a tendency to form an S shape and to be symmetrical around the inclination.

BTPS correction was done automatically by the software when the result was analyzed. $s_{Raw}$ from one run was calculated as the median value of at least five technically satisfactory loops with similar configuration and inclination.°

**Statistical Analysis**

We used analysis of variance with unbalanced block design (SAS Proc GLM) to analyze differences due to center, child, center-specific observer, accompanying observer (Pornitriva Poorisrisak), and age of the child. We included age in the analysis of center agreement because of the small number of children. Younger children could theoretically have a higher variation of $s_{Raw}$ values throughout the many visits. Our data were powered to detect a difference of 0.078 in expected log ($s_{Raw}$) between two prespecified centers. If centers are not prespecified, our data were powered to detect a difference of 0.117 using Bonferroni correction. For the normative data, we used a mixed model with repeated measurements using log-transformed $s_{Raw}$ values ad-
justed for center number. A comparison between data from center 3 with previous reported normative values was done using a two-sample \( t \) test for means with log-transformed \( s_{\text{Raw}} \) values. The calculations were done with a statistical software package (SAS, version 9.1; SAS Institute; Cary, NC).

**Results**

All seven children completed measurements at each of the six centers. The children were between 4.9 and 6.6 years old (three boys). None of the children had a history of asthma or allergy. Three children had had atopic dermatitis, three had parental atopy, and none had smoking parents.

Lung function measurements differed significantly between centers (Fig 1). \( s_{\text{Raw}} \) at centers 1 and 2 were significantly lower in all children compared with the other four centers, and center 6 had significantly higher \( s_{\text{Raw}} \) values than the other centers. Mean \( s_{\text{Raw}} \) for all six centers was 0.88 kPa/s (SD, 0.23). The within-subject SD was 0.01, and the between-center SD for each child was 0.02. Observer and age of the child did not significantly affect the measurements (\( p > 0.5 \)). For the individual results of the seven children for all six centers, see supplementary Fig 1 online.

A technician from the company (Cardinal Health) was sent to identify the problems in the deviating centers (centers 1, 2, and 6). This revealed an incorrect setting of the “ASC (automatischer schleifen-computer) Compensation” at center 1. “Time delay for compensation” was set to 20 ms and should have been 50, which resulted in 19–32% lower values. This was a factory setting not accessible to the operator. The technician found no reason for the deviating measurements at the other centers (numbers 2 and 6). It was not possible to reanalyze the data because the software saved the \( s_{\text{Raw}} \) values after the primary calculation of \( s_{\text{Raw}} \).

Subsequently, 105 preschool children (52 boys) were measured in five of the centers (Table 1) with a mean age of 5.1 years (interquartile range, 4.3 to 6.0). One child was of Latin American descent and two of Arabic descent. Centers 1 through 5 provided data for the healthy cohort. The center numbers in the biological control study represent the same center numbers in the normative study.

Mean \( s_{\text{Raw}} \) was 1.21 kPa/s (SD, 0.33) independent of height (Fig 2), weight, age, and gender (Table 1) (\( p > 0.05 \) for all estimates); within-subject SD was 0.07. There was no significant effect of center. Furthermore, there was no effect of the child’s history of atopy, parental atopy, or smoking (\( p > 0.05 \) for all estimates).

Center 3 used the exact same equipment (hardware and software) as in our previous report on normative data. A comparison was made to ensure that time (10 years between the two studies) did not

### Table 1—Normative Study: Clinical Characteristics of the Children From the Five Centers*

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Center No.</th>
<th>Estimate (95% CI)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children, no.</td>
<td>21</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td>( s_{\text{Raw}} )</td>
<td>1.30 (0.32)</td>
<td>1.09 (0.24)</td>
<td>1.26 (0.31)</td>
</tr>
<tr>
<td>Age,† yr</td>
<td>5.39 (1.16)</td>
<td>5.24 (1.04)</td>
<td>5.07 (1.18)</td>
</tr>
<tr>
<td>Male gender‡</td>
<td>11 (52)</td>
<td>15 (54)</td>
<td>16 (55)</td>
</tr>
<tr>
<td>Weight,‡ kg</td>
<td>22.0 (4.29)</td>
<td>20.0 (4.84)</td>
<td>20.4 (4.32)</td>
</tr>
<tr>
<td>Height,‡ cm</td>
<td>115.2 (8.95)</td>
<td>110.2 (9.90)</td>
<td>112.3 (9.79)</td>
</tr>
<tr>
<td>Rhinitis‡</td>
<td>4 (19)</td>
<td>0 (0)</td>
<td>4 (14)</td>
</tr>
<tr>
<td>Dermatitis‡</td>
<td>2 (9.5)</td>
<td>10 (36)</td>
<td>5 (17)</td>
</tr>
<tr>
<td>Parental‡ atopy</td>
<td>12 (57)</td>
<td>13 (46)</td>
<td>21 (75)</td>
</tr>
<tr>
<td>Smoking‡</td>
<td>4 (19)</td>
<td>8 (29)</td>
<td>4 (14)</td>
</tr>
</tbody>
</table>

*CI = confidence interval.

†Values are given as the mean (SD).

‡Values are given as No. (%).

Figure 1. Center agreement study: least squares mean for \( s_{\text{Raw}} \) for six centers with 95% confidence interval (CI) [software versions are indicated in parentheses].
have an effect on sRaw measurements before we pooled the data. We previously reported normative values from a population of 121 children, 2 to 5 years of age; the mean sRaw was 1.31 (SD, 0.20). The center 3 result was as follows: mean sRaw, 1.26; SD, 0.31. We compared the two data sets with a two-sample \( t \) test for means with log-transformed sRaw values (\( p \) value 0.20). There was no significant difference between these previous data and the current normative data. Therefore, we pooled the previous data (121 children) and the current normative data (105 children) (Fig 3).

**DISCUSSION**

**Between-Center Variation**

sRaw offers a method for clinical monitoring and research during the critical period of growth and development early in life. The method is feasible from the age of 2 years, and the precision is high.\(^1\)-\(^3\) However, the present study showed that the accuracy of sRaw measurements in young children was flawed due to errors in the factory setting in one center. It is the key message of our study that center effects were seen and could only be explained by differences in the software hidden from the end user. After correcting the factory settings at the deviating center, there were no longer differences between the centers, and normative values were generated in this multicenter setting. The problem was not discovered by the standard calibration of flow, box leak, and internal box pressure. Current calibration only assesses flow measured by the pneumotachograph, leak from the box, and pressure transducer. The available calibration does not assess the final resistance measure, which is generated by algorithms buried in the software with settings often inaccessible to the end user. Thus, errors in software or mechanics may go unnoticed with a potential impact on clinical evaluation and flawed accuracy as illustrated in our study. A mechanical infant lung model analog has previously been developed for quality control of infant whole-body plethysmographs,\(^7\) but a model testing for preschool children is not available to the end user. This study suggests the need for development of methods for control of the actual resistance measure for young children and not only the flow and box leakage. Without such a proof of accuracy, normative values generated at other centers may not be applicable. Until a mechanical standard becomes available, the biological standard (healthy subjects) is the only possible substitute.

We used a standardized protocol including standard calibration of flow, box leakage, and internal box pressure in six Danish centers at secondary and tertiary referral hospital departments. The six centers included in the study of accuracy were spread over the country, which prevented measurements on the same day. Therefore, the day-to-day variability reduced the sensitivity by which we could identify outliers among the centers. The children were not trained before entering the study. The visit order was randomized to ensure a possible difference between the first and second visit did not bias the center variation.

In the current study, the within-subject SD on the same day and center was 0.01, and the within-subject SD between centers was 0.02. In our previous study, the precision (repeatability) of sRaw measurements 9 days (mean) apart in young children with asthma (asymptomatic during the study period) was found to have an intraclass coefficient of 0.87 (within-subject SD 0.03) for baseline measurements between occasions.\(^10\) The higher within-subject SD could be explained by the asthma status of children in the previous study.

The current study was designed to find a possible center effect. We were able to account for any possible observer bias by having a center-specific

---

**Figure 2.** Normative study: healthy sRaw data against height for five centers.

**Figure 3.** Normative data from the multicenter study (five centers) and the previous study by Klug and Bisgaard.\(^3\)
Normative Data

In the second part of the study, a center effect could not be found though, probably because center 1 was corrected and center 6 did not participate. Atopy and smoking did not significantly differ between centers (Table 1). The high incidence of parental atopy could be a selection bias, but we did not find a statistical difference between atopic disposed and nondisposed children.

We previously reported normative values from a population of 121 children 2 to 5 years of age\(^3\); the mean sRaw was 1.31 (SD, 0.20). The previous study differs from the current in several of the following aspects: (1) measurements were made at one center; (2) children exposed to tobacco smoke and anyone with a history of eczema or doctor-diagnosed atopy in first-degree relatives were excluded from the study; and (3) the study included more 2- and 3-year-old children. Many of these measurements had an accompanying adult in the whole-body plethysmograph. The current data included only children who had performed a lung function measurement alone. There was no difference between sRaw measurements in the previous normative and the current normative data for the same center using the very same equipment. Therefore, we decided to pool the two sets of normative data, showing the normal sRaw in young children to be 1.27 kPa/s (SD, 0.25 kPa/s) independent of age, height, and gender (Fig 3).

In conclusion, using a biological control we revealed errors in the accuracy of sRaw measurements at some centers despite normal calibration of the mechanical components. This study highlights the need for the development of equipment allowing control of the actual resistance measurements and not only some of its mechanical components. Until such equipment becomes available, the only option is to use healthy subjects to assure that the absolute values measured are similar to the normative values reported in this and a previous report of healthy young children.

ACKNOWLEDGMENTS: We thank all the children and parents who participated in the study. We also thank S. Pedersen, B. Christensen, and T. Duns (Danish Pediatric Asthma Center, Kolding); H. Andreassen and A. K. Thomsen (Pediatric Department, Naestved Hospital); L. Johnsen, H. Niekrenz, and P. O. Schioetz (Danish Pediatric Asthma Center, Aarhus University Hospital, Skejby); K. Mathiesen, L. Klaassoe, S. V. Thorsen, S. M. Jensen, and C. B. Piper (Danish Pediatric Asthma Center, Copenhagen University Hospital, Gentofte); and C. Petersen (Intramedic) for advice and practical help; and Jensens Boefhus for sponsoring a dinner for the families.

REFERENCES

Accuracy of Whole-Body Plethysmography Requires Biological Calibration
Porntiva Poorisrisak, Carsten Vrang, Jorn Molgaard Henriksen, Bent Klug, Birgitte Hanel and Hans Bisgaard
Chest 2009;135; 1476-1480
DOI 10.1378/chest.08-1555

This information is current as of February 14, 2011

Updated Information & Services
Updated Information and services can be found at:
http://chestjournal.chestpubs.org/content/135/6/1476.full.html

References
This article cites 10 articles, 7 of which can be accessed free at:
http://chestjournal.chestpubs.org/content/135/6/1476.full.html#ref-list-1

Cited By
This article has been cited by 3 HighWire-hosted articles:
http://chestjournal.chestpubs.org/content/135/6/1476.full.html#related-urls

Permissions & Licensing
Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
http://www.chestpubs.org/site/misc/reprints.xhtml

Reprints
Information about ordering reprints can be found online:
http://www.chestpubs.org/site/misc/reprints.xhtml

Citation Alerts
Receive free e-mail alerts when new articles cite this article. To sign up, select the “Services” link to the right of the online article.

Images in PowerPoint format
Figures that appear in CHEST articles can be downloaded for teaching purposes in PowerPoint slide format. See any online figure for directions.