Feasibility and Variability of Neonatal and Infant Lung Function Measurement Using the Single Occlusion Technique*

Nienke Katier, MD; Cuno S. P. M. Uiterwaal, MD, PhD; Brita M. de Jong, MD; Jan L. L. Kimpen, MD, PhD; and Cornelis K. van der Ent, MD, PhD

Introduction: For possible use as a predictor of wheezing illnesses in routine care, we evaluated the feasibility and variability of measurement of passive respiratory mechanics in a large, open population of healthy neonates and infants.

Methods: As part of the ongoing Wheezing Illnesses Study Leidsche Rijn, respiratory compliance (Crs), respiratory resistance (Rrs), and time constant (τrs) were measured during natural sleep in 450 healthy term neonates and infants using the single-occlusion technique (SOT). Interobserver and intraobserver variability of data sampling and the subsequent selection and analysis of occlusions as well as intrameasurement variability were examined.

Results: Technically acceptable lung function measurements could be performed in 328 infants (73%). Low intraobserver and interobserver variability was found for both data sampling (intraclass correlation coefficient [ICC] ≥ 0.87) and for selection and analysis of occlusions (ICC ≥ 0.99). Intrameasurement variability was low, with a mean intrameasurement coefficients of variation for Crs, Rrs, and τrs of 8.5%, 10.4%, and 15.4%, respectively. Averaging three or more occlusions resulted in stable values of Crs, Rrs, and τrs.

Conclusion: Results of this study indicate that feasibility and variability of lung function testing using the SOT is acceptable for use in large populations of healthy neonates and infants in routine care.

(CHEST 2005; 128:1822–1829)

Key words: birth cohort; infant; passive respiratory mechanics; respiratory function test; variability

Abbreviations: ATS = American Thoracic Society; CI = confidence interval; Crs = respiratory compliance; CV = coefficient of variation; ERS = European Respiratory Society; ICC = intraclass correlation coefficient; Rrs = respiratory resistance; SOT = single occlusion technique; τrs = time constant; Vplat = volume at pressure plateau; WHISTLER = Wheezing Illnesses Study Leidsche Rijn

Lower respiratory tract illnesses are a major cause of morbidity and mortality in infancy and childhood, with large impact on health care.1,2 Epidemiologic research concerning the development and prediction of childhood respiratory disease may benefit from lung function measurements in newborn infants. Only a few small prospective cohort studies3–7 have investigated premorbid lung function in association with subsequent respiratory disease; these studies used different lung function techniques and showed conflicting results. The major limitation is the lack of simple, reliable, and reproducible lung function tests that are applicable in a large, open population of healthy infants. Although several infant lung function techniques have been developed.

*From the Departments of Pediatric Pulmonology (Drs. Katier, de Jong, and van der Ent), Pediatric Infectious Disease (Dr. Kimpen), and Julius Center for Health Sciences and Primary Care (Dr. Uiterwaal), University Medical Center Utrecht, Utrecht, the Netherlands.

This study was funded by the Health Research and Development Council of the Netherlands grant No. 2001–1-1322, and by an unrestricted grant from Glaxo Smith Kline Netherlands.

Manuscript received December 2, 2004; revision accepted February 8, 2005.

Reproduction of this article is prohibited without written permission from the American College of Chest Physicians (www.chestjournal.org/misc/reprints.shtml).

Correspondence to: Cornelis K. van der Ent, MD, PhD, Department of Pediatric Pulmonology, University Medical Center Utrecht, Internal Postal Code: KH 01.419.0, PO Box 85090, 3508 AB Utrecht, the Netherlands; e-mail: k.vanderent@wkz.azu.nl
throughout the last 25 years with considerable methodologic progress and refinement, most require sophisticated equipment and sedation and, therefore, application is limited to specialized research laboratories.

The single occlusion technique (SOT) is a lung function technique for assessment of passive respiratory mechanics. First proposed as a simple method for measuring respiratory compliance \((C_{rs})\) by Ohlsnky et al. and later refined by Lesouef et al. and Mortola et al., this technique measures \(C_{rs}\), respiratory resistance \((R_{rs})\), and time constant \((\tau_{rs})\) of the total respiratory system. It is a noninvasive, easily applicable, and time-saving technique for assessment of lung function in spontaneously breathing infants and young children. Guidelines regarding equipment, signal processing, data handling, and control and acceptance criteria have been published by the European Respiratory Society (ERS)/American Thoracic Society (ATS) Task Force on standards for infant respiratory function testing. Measurement during natural sleep and in awake infants is possible, which makes this technique particularly suitable for infant lung function testing. In order to be used as a predictor of wheezing illnesses in routine care, information on the feasibility and variability of the SOT is necessary. However, little is known about the feasibility and variability of the SOT when applied in a large population of healthy neonates and infants in routine care.

The aim of this study was to examine the feasibility and variability of \(C_{rs}\), \(R_{rs}\), and \(\tau_{rs}\) measurements using the SOT in a large, population-based birth cohort in a routine care setting. The following questions were addressed: (1) is it feasible to perform SOT measurements in a large group of healthy neonates and infants, (2) what is the interobserver and intraobserver variability of data sampling and subsequent selection and analysis of occlusions, and (3) what is the intrameasurement variability of \(C_{rs}\), \(R_{rs}\), and \(\tau_{rs}\) values?

**Materials and Methods**

**Study Population**

All infants were recruited as part of the ongoing Wheezing Illnesses Study Leidsche Rijn (WHISTLER). Briefly, WHISTLER is a single-center, prospective, population-based birth cohort study on determinants for wheezing illnesses (including neonatal lung function). WHISTLER ultimately aims to derive a comprehensive risk score for wheezing illnesses appropriate for use in primary health care as described previously. Lung function measurements were performed in healthy neonates and infants <2 months old before any respiratory illness was present. Exclusion criteria were gestational age <36 weeks, major congenital abnormalities, and neonatal respiratory disease. Approval of the pediatric medical ethics committee of the University Medical Center Utrecht was given for this study, and written informed parental consent was obtained.

**Recording Equipment**

Airflow was measured using a heated Lilly-type pneumotachometer (series 8300, linear range, 0 to 10 L/min; Hans Rudolph, Kansas City, MO) attached to a face mask (infant mask, size neonate; Hans Rudolph). The mask was sealed to the infant’s face using therapeutic silicone putty (Magic Putty, medium; Oldelft Benelux BV; Delft, the Netherlands) to prevent air leaks and to minimize dead space. Pressure changes at the airway opening were measured with a pressure transducer (Type 163PC01D75; Honeywell, Morristown, NJ). Volume was obtained by electronic integration of the airflow signal. Flow, volume, and pressure were digitized with a sampling rate of 200 Hz and interfaced to a computer for real-time display, storage, and analysis. Before each measurement, calibration of flow and volume signals was performed using a 100-mL precision syringe (ViaSys Health; Höchberg, Germany). The pressure transducer was calibrated over the expected range using a pressure transducer tester (Verical; Utah Medical Products; Midvale, UT). Lung function data were calculated offline using a custom-built software package (Luna 1.6; University Medical Center Utrecht; Utrecht, the Netherlands). Parameters displayed during analysis are as follows: maximal volume and volume at pressure plateau (\(V_{plat}\)); maximum volume and \(V_{plat}\) in percentage; the SD of \(V_{plat}\) and pressure plateau, the regression coefficient \((r)\) of the expiratory part of the flow-volume curve, volume at intercept, and \(C_{rs}\), \(R_{rs}\), and \(\tau_{rs}\). Regression length can be varied manually between 40% and 80% of expired volume. Occlusions can be accepted or disregarded based on the displayed values. Final results (all trials or summary statistics) can be exported to standard spreadsheet software.

**Measurement Protocol**

Measurements of passive respiratory mechanics \((C_{rs}, R_{rs}, \text{and } \tau_{rs})\) were performed without the use of any sedation in a routine care setting. Lung function measurements were performed after feeding in supine position with a neutral head/neck posture. Data collection was confined to consecutive periods of quiet sleep in which posture was stable and respiration was regular. Once each infant was settled in quiet sleep, the mask was placed gently over the nose and mouth, creating a continuous seal. Brief occlusions were performed manually at maximal inspiration. The occluded airway opening was released when a plateau in mouth pressure was reached. The infant’s breathing was allowed to stabilize between each occlusion for a minimum of five breaths. Occlusions were considered acceptable when they met the criteria of the ERS/ATS Task Force: (1) smooth expiration within 10% of the previous expiration; (2) expiration without evidence of glottic closure, braking, or active expiratory effort; (3) duration of a pressure plateau ≥ 100 ms; (4) variability of pressure plateau < 0.01 kPa; and (5) linearity of descending part of the passive flow-volume loop over at least 40% of expiration with \(r^2 > 0.99\). At least three technically acceptable occlusions (up to 31) were used to calculate mean \(C_{rs}\), \(R_{rs}\), and \(\tau_{rs}\).

Several levels of the process of data sampling and the subsequent selection and analysis of the occlusions can cause variability in the ultimate values of \(C_{rs}\), \(R_{rs}\), and \(\tau_{rs}\). First, differences in infant and device handling might result in differences in the infant’s respiration between different observers (interobserver variability of data sampling). Second, repeated data sampling by the same observer in the same child might result in differences in
Feasibility in unselected infants (n=450)

Data sampling

Selection and analysis of occlusions

Intra-measurement variability of Crs, Rs, and τrs values (n=50)

**Figure 1.** Study overview.

due to small changes, for example, in the child's position of the face mask (intraobserver variability of data sampling). Thirdly, after sampling and electronic storage of measurement tracings, differences in selection and analysis of occlusions from the recorded tracings might result in different values of Crs, Rs, and τrs between observers (interobserver variability of selection and analysis of occlusions). Fourthly, this process of selecting and analyzing occlusions might also cause differences in results within one observer (intraobserver variability of selection and analysis of occlusions). Ultimately, the number of occlusion used for the calculation of Crs, Rs, and τrs within one measurement and analyzed by one observer might influence the ultimate values of Crs, Rs, and τrs (intrameasurement variability). Figure 1 gives an overview of the study components to address these issues. For the interobserver variability study of data sampling, 40 infants underwent repetitive measurements within a short time interval by two different observers. For the intraobserver variability study of data sampling, another 40 infants underwent repetitive measurements performed by one of the observers. For the evaluation of interobserver variability of selection and analysis of occlusions, two different observers interpreted 259 single lung function measurements. Observers independently selected and analyzed occlusions based on the criteria of the ERS/ATS Task Force. To evaluate intraobserver variability of selection and analysis of occlusions, one observer interpreted 40 lung function measurements twice blinded to the results earlier obtained. Thirty measurements were randomly selected in order to analyze the intrain measurement variability of average Crs, Rs, and τrs values based on 2 to 10 sequential occlusions. To obtain a graphic representation of the difference between and within observers, differences were plotted against the mean and the mean difference, the limits of the agreement, and the 95% confidence intervals (CI)'s for the limits were calculated, as described by Bland and Altman. Intrameasurement variability was also expressed as coefficient of variation (CV), i.e., the ratio of the SD to the mean. Differences were considered statistically significant if the p value was ≤ 0.05. Statistical analysis was performed using the statistical software (SPSS version 11.0, SPSS; Chicago, IL).

### Results

#### Feasibility

Subject characteristics and the results of the pulmonary function tests are shown in Table 1. Of the 450 infants recruited for this study, technically acceptable lung function measurements could be performed in 311 infants (69%) on the first occasion. Twenty-three parents were willing to return for a second visit, resulting in additional successful measurements in 17 infants (4%). Failure of technically acceptable measurements was mainly due to failure to fall asleep naturally within 1.5 h of study onset (16%), consistent failure to relax during occlusions (5%), marked expiratory airflow braking (4%), and persistent leak around the face mask (2%). The median number of successful occlusions per measurement was 5 (range, 3 to 31). Eighty-two percent of these infants were white. A history of active maternal smoking was positive in 8.9% of infants, and postnatal environmental smoking by parents or caregivers was present in 4.5%. The characteristics of infants with unsuccessful lung function measurements were not different from infants with successful measurements, except for a significantly younger age (Table 1). This is also shown in Figure 2, top, A, as failure rates decline with increasing age at mea-

### Table 1—Patient Characteristics and Results From Respiratory Function Measurements in 450 Healthy Infants*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total Group</th>
<th>Successful Measurement</th>
<th>Unsuccessful Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infants, No.</td>
<td>450</td>
<td>328</td>
<td>122</td>
</tr>
<tr>
<td>Male/female gender, No.</td>
<td>209/241</td>
<td>152/176</td>
<td>57/65</td>
</tr>
<tr>
<td>Age, wk</td>
<td>4.6 ± 1.3</td>
<td>4.7 ± 1.3</td>
<td>4.3 ± 1.1†</td>
</tr>
<tr>
<td>Height, cm</td>
<td>54.3 ± 3.6</td>
<td>54.3 ± 2.5</td>
<td>54.3 ± 2.7</td>
</tr>
<tr>
<td>Body weight, kg</td>
<td>4.31 ± 0.67</td>
<td>4.27 ± 0.69</td>
<td>4.40 ± 0.62</td>
</tr>
<tr>
<td>Crs, mL/LPa</td>
<td>40.67 ± 11.18</td>
<td>9.58 ± 2.53</td>
<td></td>
</tr>
<tr>
<td>Rs, mL/LPa/kg</td>
<td>7.95 ± 2.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>τrs, s</td>
<td>0.321 ± 0.132</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Data are presented as mean ± SD unless otherwise indicated.
†Successful vs unsuccessful measurement (p = 0.002).
measurement (range of age at measurement, 1.4 to 9.1 weeks). Failure rates were also dependent on the experience of the observers (Fig 2, bottom, B). We divided the 450 measurements in three parts of 150 measurements numbered in sequence of time. The failure rate during the first 150 measurements was 27%, during the last 150 measurements it decreased to 15%. Age was not a confounding variable; ie, the last 150 measurements were not made in older children.

Variability of Data Sampling

For the intraobserver variability study of data sampling, the ICCs were calculated for Crs, Rrs, and $\tau_{rs}$, and were found to be 0.96, 0.98, and 0.95, respectively. Between two different observers, the ICCs for Crs, Rrs, and $\tau_{rs}$ were 0.97, 0.91, and 0.87, respectively. The mean differences of Crs, Rrs, and $\tau_{rs}$ between two consecutive measurements performed by the same observer and two different observers, the limits of agreement, and the 95% CIs of the limits, as described by Bland and Altman,17 are shown in Table 2.

Variability of Selection and Analysis of Occlusions

For the intraobserver variability study of selection and analysis of occlusions, ICCs were calculated for Crs, Rrs, and $\tau_{rs}$, and were found to be 0.99, 0.99, and 0.99, respectively. Between two different observers, the ICCs for Crs, Rrs, and $\tau_{rs}$ were again 0.99, 0.99, and 0.99, respectively. The mean differences of Crs, Rrs, and $\tau_{rs}$ between and within interpreters of lung function measurements, the limits of agreement and the 95% CIs of the limits are shown in Table 2.
Data sampling
Values
Intrameasurement Variability of Crs, Rrs, and τrs

For 30 randomly selected measurements, the mean Crs, Rrs, and τrs and their SDs of 2 to 10 occlusions within one measurement were calculated. Intrameasurement CVs ranged from 5.1 to 12.0% with a mean of 8.5% for Crs; from 4.5 to 14.8% with a mean of 10.4% for Rrs; and from 10.2 to 20.3% with a mean of 15.4% for τrs. The CVs of Crs, Rrs, and τrs within one measurement correlated with the mean Crs (r = 0.57, p = 0.001), Rrs (r = 0.80, p < 0.001), and τrs (r = 0.85, p < 0.001) within one subject; ie, the SD was proportional to the mean.

Another way to investigate intrameasurement variability is to evaluate the intrameasurement stability of average Crs, Rrs, and τrs values. ICCs were calculated for 2 to 10 occlusions within one measurement, as shown in Figure 3. The nearer the ICC is to 1, the more stable the average ratio. The use of two occlusions would result in a value of Crs with an ICC of 0.95, a value of Rrs with an ICC of 0.91, and a value of τrs with an ICC of 0.84. An increase in the number of occlusions increases the ICC of the subsequently measured values of Crs, Rrs, and τrs. Averaging three or more occlusions resulted in stable values (ICC ≥ 0.90) for all three parameters; averaging more than six curves did not improve the stability of the mean values.

### Discussion

Using a large and unique data set of neonatal lung function measurements, this study was able to assess the feasibility and variability of the SOT in a large, open population of healthy neonates and infants. In the majority of infants, technically acceptable lung function measurements could be performed, with intrameasurement CVs ranging from 8 to 15%. Low interobserver and intraobserver variability was found for data sampling and for the subsequent selection and analysis of occlusions. Results of this study indicate that feasibility and variability of lung function testing using the SOT is acceptable for use in open populations of healthy neonates and infants and may be used as a possible predictor of wheezing illnesses in routine care.

Although feasibility and variability of the SOT are described in a routine care setting, it has to be considered that measurements are still performed within a research design with dedicated and thorough highly experienced researchers. Application of the SOT by different observers in routine care may in practice turn out to be less accurate.

### Feasibility

In the present study, measurements were performed without use of sedation in a routine care setting by trained personnel and, as strictly as possible, according to the criteria of the ERS/ATS Task Force.14 Relatively few studies have reported failure rates when describing occlusion techniques.18,19 Failure rates in a setting as described in the present study are to our knowledge not available. Therefore, it is difficult to ascertain whether the failure rates in the present studies are attributable to the strict criteria applied or to the type of setting and study population. Successful lung function testing was
Figure 3. ICCs of (top, A) Crs calculated as an average of the number of sequential occlusions; (center, B) Rrs calculated as an average of the number of sequential occlusions; and (bottom, C) τrs calculated as an average of the number of sequential occlusions.
feasible in 73% of the infants. The main reason for measurement failure was not falling asleep naturally within 1.5 h of study onset (16%). The use of sedation would diminish these failure rates; however, this would limit the technique to specialized research laboratories and would be more time consuming and invasive. Five percent of infants consistently failed to relax during occlusions generally recognized as a rising pressure at airway opening when the infants made an expiratory effort against the occlusion. Stocks et al.\textsuperscript{20} reported a percentage of 9.5% in sedated and anesthetized infants during the first 2 years of life. Marked expiratory braking was the reason for 4% of failures. Potential errors in assessing passive respiratory mechanics due to expiratory airflow modulation could be detected in the present study by a linearity of the flow-volume curve ($r^2 < 0.99$). Problems in the use of the SOT due to expiratory braking have been reported previously in approximately 10 to 30% of infants.\textsuperscript{19,21} One explanation that could account for the differences in failure rates between the present study and others is the infant’s age at measurement. The few previous studies reporting failure rates measured infants that were either older\textsuperscript{20,21} or preterm.\textsuperscript{19} Another explanation could be the experience of observers performing the measurements. Previous studies\textsuperscript{18,19} reporting failure rates measured only small groups of infants. In this study, a large group of infants was measured, with failure rates declining with increasing experience of the observers. Failure rates during the first 150 measurements were, however, lower than during the second 150 measurements. This may be due to the fact that during the first measurements, an experienced supervisor was always present to take over the measurement when necessary. During the second and last 150 measurements, observers started measuring without supervision; as experience increased, failure rates declined.

The values of $Rrs$ are in close agreement with Hanrahan et al.\textsuperscript{22} but are higher than those reported by Masters et al.\textsuperscript{23} Lodrup Carlsen et al.\textsuperscript{24} and Keiterer and Muller.\textsuperscript{25} $Crs$ levels in the current study are somewhat lower than reported in previous studies, but comparable to values reported by Stocks et al.\textsuperscript{26} and Migdal et al.\textsuperscript{27} One explanation that could account for the differences is that the postnatal age in the population assessed in the present study was higher, except for the study of Masters et al.\textsuperscript{23} However, in this study only seven healthy infants were assessed. Differences may also be due to use of different recording equipment in different studies. The published recommendations and guidelines regarding equipment\textsuperscript{12} and signal processing and data handling\textsuperscript{13} may improve comparison of passive respiratory mechanics measurements of different studies and avoid misinterpretation of results.

**Variability of $Crs$, $Rrs$, and $\tau rs$ Measurements**

Limited data are available on the variability of the SOT. Data on interobserver and intraobserver variability of the SOT data sampling are to our knowledge not available. Passive respiratory mechanics are frequency dependent and may be influenced by upper airway resistance (laryngeal obstruction) to flow. Dynamic increase in functional residual capacity often occurs in neonates and might also cause difference in results of repeated measurements within and between different observers. However, we found high agreement within and between observers on repetitive measurements of neonatal lung function within a short time span. Obviously, these disturbing factors do not importantly limit the use of the SOT in the open population setting.

The variability of the method is also highly dependent on selection and analysis of occlusions that are used for calculating $Crs$, $Rrs$, and $\tau rs$. An excellent level of agreement was found between two observers with acceptably narrow limits of agreement. The intraobserver agreement was equally high. The low interobserver and intraobserver variability of data selection and analysis may be due to the strict use of the criteria of the ERS/ATS task force. The use of international criteria for selection and analysis of occlusions may greatly decrease the variability of the method.

A few reports on intramasurement variability of lung function measurements using the SOT exist. The intramasurement variability of lung function measurements consisting of 2 to 10 occlusions in the present study was low compared to earlier reports of Lodrup-Carlsen et al.\textsuperscript{24} and Hanrahan et al.\textsuperscript{22} This difference may be due to the large sample size of the present study, with increasing experience of the observers and the fact that measurements were confined to periods of quiet sleep. ICCs for statistical evaluation of the stability of $Crs$, $Rrs$, and $\tau rs$ after 2 to 10 occlusions within one measurement showed that after three occlusions reliable values can be found, indicating that a minimum of three occlusions, as recommended, is acceptable.\textsuperscript{14} The ICCs after six occlusions rose to 0.97, 0.98, and 0.97 for $Crs$, $Rrs$, and $\tau rs$, respectively, with no improvement of stability when averaging more occlusions.

**Conclusion**

This study has demonstrated that it is feasible to measure lung function using the SOT in unsedated healthy neonates and infants in routine care with low interobserver and intraobserver variability of data
sampling and the subsequent selection and analysis of occlusions. The results indicate that the SOT is a simple, accurate, and reliable technique to assess lung function in large open populations of healthy neonates and infants. Further investigations are needed to explore the physiologic determinants of passive respiratory mechanics in healthy neonates, prenatal and early life events that influence lung function in infancy, and the uses and limitations in epidemiologic research, e.g., in the evaluation of the possible predictive value of neonatal lung function for later respiratory disease.

ACKNOWLEDGMENT: The authors thank Claire Nollen and Cora de Kiviet for assistance during lung function measurements and Myriam Olling-de Kok for secretarial work.

REFERENCES