A new non-invasive method of infant spirometry demonstrates a level of repeatability that is comparable to traditional methods

Mariani H. L. Bentsen (Mariann.Haavik.Lysfjord.Bentsen@helse-bergen.no, Mariann.Bentsen@k2.uib.no)1,2, Øystein A. Haaland1,3, Lise B. Lønning1, Hrefna K. Gudmundsdottir1, Trond Markestad1,2, Thomas Halvorsen1,2

1.Department of Pediatrics, Haukeland University Hospital, Bergen, Norway
2.Department of Clinical Science, Faculty of Medicine and Odontology, University of Bergen, Bergen, Norway
3.Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway

Keywords
Infant, Plethysmography, Premature, Respiratory function tests, Spirometry

Correspondence
M Bentsen, MD, Department of Pediatrics, Haukeland University Hospital, N-5021 Bergen, Norway.
Tel: +47 55970871 | Fax: +47 55975147 | Emails: Mariann.Haavik.Lysfjord.Bentsen@helse-bergen.no, Mariann.Bentsen@k2.uib.no

Received
11 December 2014; revised 9 April 2015; accepted 13 August 2015.
DOI:10.1111/apa.13155

ABSTRACT

Aim: The FloRight system provides novel non-invasive infant spirometry based on electromagnetic inductance plethysmography. We investigated the consistency of repeated measurements carried out in a Norwegian neonatal intensive care unit (NICU) using the system and how well these were tolerated.

Methods: Tidal flow-volume loops were obtained from 10 preterm infants at discharge, 10 stable growing preterm infants weighing about 1500 g and 10 term-born infants. A nurse experienced with the system measured all patients before and after meals, and these measurements were repeated by nurses new to the system.

Results: The measurements were well tolerated by the infants. The repeatability for the two parameters 'tidal volume' (Vt) and 'time to peak tidal expiratory flow to total expiratory time' (Tptef/Te) were relatively poor, similar to previous methods. However, the repeatability was good for the new 'flow-volume gravity mid-point' (FVg) parameter. Repeatability was better for term than preterm infants, when measurements were obtained by the experienced nurse and for measurements carried out before meals.

Conclusion: The FloRight system proved feasible in a NICU setting. The repeatability of the lung function measurements was similar to those reported for traditional infant spirometry. The nurse’s experience and the relationship to meals appeared to be important.

INTRODUCTION

Lung diseases are common in neonates and early childhood. However, our understanding of the detailed lung mechanics in this age group is limited, during both health and disease, as methods for lung function measurements are complex and not readily available (1). This situation hampers our approach to medical care and also complicates lifelong tracking of lung function. Infant spirometry most commonly involves measuring airflow through a face mask, but other methods have also been used, such as whole-body plethysmography or respiratory inductance plethysmography (1–5). Most of these methods are time-consuming, require some form of sedation and disturb the infants, which means that they can only be used for brief periods and are not really suitable for clinical settings other than dedicated specialist laboratories. In addition, a mask may alter the pattern of breathing and the added dead space will influence the parameters measured (6–8).

The FloRight system (VoluSense, Bergen, Norway) is a novel method of infant spirometry, based on electromagnetic inductance plethysmography. The method is non-invasive, easy to calibrate, does not require sedation and is free of dead space artefacts.

Clinical validation studies comparing the FloRight system with ultrasonic flow meters (9,10) or pneumotachographs (11) have shown somewhat conflicting results for tidal volume measurements. Petrus et al. found that tidal volumes were, on average, 1.3 mL/kg lower with the FloRight system (10), while others found reasonable agreement between the FloRight system and mask-based methods (9,11). In all of these studies, the other tidal breathing parameters corresponded with each other and all the

Key notes
- The FloRight system is a novel non-invasive method for infant spirometry based on electromagnetic inductance plethysmography.
- We investigated the consistency of repeated measurements carried out on full and preterm infants by different nurses before and after meals and found that FloRight compared well with results from more intrusive methods.
- Repeatability was better for term infants, when measurements were obtained by a more experienced nurse and when carried out before meals.
systems were able to discriminate between healthy and diseased lungs.

The aim of this study was to assess the clinical applicability of the FloRight system, by examining the consistency of repeated measurements and how well the measurements were tolerated by infants.

**PATIENTS AND METHODS**

**Subjects**

We studied 30 infants in the neonatal intensive care unit (NICU) at Haukeland University Hospital, Bergen, Norway, who did not require assisted ventilation or oxygen supplementation at the time of study. These comprised 10 preterm infants examined at discharge and 10 stable growing preterm infants weighing approximately 1500 g, who were basically healthy, but still needed help to eat and regulate their temperature and were under surveillance for apnoeas. Six of the preterm infants examined at 1500 g and five of those examined at discharge had respiratory distress shortly after birth. The other ten were full-term infants who were in the NICU because of initial hypoglycaemia, anaemia, thrombocytopenia or suspected infection, but were diagnosed as healthy after short-term observation and, or, treatment. None of them had respiratory symptoms. The characteristics of the participants are shown in Table 1.

The study was approved by the Regional Committee on Medical Research Ethics of Western Norway and reported to ClinicalTrials.com (ID no. NCT01057472). Informed, written consent was obtained from the parents.

**Equipment**

The FloRight system consists of two thin metal wire coils that are sewn in a circular fashion into a thin, soft cotton vest, which is worn over the entire chest and abdomen. The upper coil covers the chest and the lower coil covers the abdomen, enabling separate analyses. A weak (100 mA) alternating electric current is passed through the coils, producing a magnetic field around the child’s torso. The strength of this field is measured by a sensor positioned directly above the cot. The system is calibrated against a volume reference cylinder in the cot at the side of the infant, so that the system knows the exact volume contained in the vest. This volume expands and decreases in proportion to the infant’s breathing movements. The volume signals are converted into flow signals by numerical differentiation, and flow-volume loops and tidal breathing parameters are calculated.

**Measurements**

We selected the correct-sized vest by measuring the infant’s chest circumference at the level of the nipple. It was put over a thin layer of clothing and covered the infant from the armpits to the groin. The vest was tightened so that it provided an adequate fit but did not constrict the torso or inhibit the infant’s breathing. All measurements were carried out at the cot with the infant in a supine position. The antenna was positioned straight above the bed at the height specified by the manufacturer, and before any measurements were carried out, the system was calibrated by placing a cylinder of known volume next to the child. Four measurement sessions, each lasting approximately five minutes, were carried out for each baby by two different nurses, Nurse A and Nurse B, in the 30 min before a planned meal in the following sequence: A1 before, B1 before, A2 before and B2 before. There was a five-minute interval between each of the four measurements. The measurements were then repeated twice by Nurse A approximately 10–15 min after a meal, and these were called measurements A3 after and A4 after. The vest was loosened and tightened between each measurement session, but not removed from the baby. The infants were relaxed, often sucking on a pacifier, and quietly awake or asleep during varying proportions of the measurement sessions. No sedation was used. The infant’s oxygen saturation and heart rate were monitored by pulse oximetry and the respiratory rate by the FloRight system. The infant’s behaviour, their facial expression and breathing patterns were subjectively observed and recorded by the examining nurse.

Seven nurses took part in the study, and they all received basic training in the equipment used in the study, including the manufacturer’s operating instructions. Nurse A had independent experience in performing the test before the formal study was initiated and was the same for all the infants. Nurse B varied and was one of six different nurses. However, Nurse B was always the same for the same infant, meaning that each infant was measured by the same two nurses.

**Data selection**

The traces were visually inspected to select two or more segments that consisted of a minimum of 10 consecutive stable breaths to enable us to calculate the tidal breathing parameters. The selection criteria were no obvious artefacts, no sighs and no obvious changes in the depth of breathing.

<table>
<thead>
<tr>
<th>Table 1 Characteristics of the participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term infants</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Number of infants</td>
</tr>
<tr>
<td>Gestational age at birth in weeks</td>
</tr>
<tr>
<td>Birthweight in grams</td>
</tr>
<tr>
<td>Age at test in days</td>
</tr>
<tr>
<td>Postmenstrual age at test in weeks</td>
</tr>
<tr>
<td>Weight at test in grams</td>
</tr>
<tr>
<td>Vt (mL/kg)</td>
</tr>
<tr>
<td>Tptef/Te</td>
</tr>
<tr>
<td>FVg</td>
</tr>
</tbody>
</table>

Vt, tidal volume; Tptef/Te, time to peak expiratory flow as a ratio of total expiratory time; FVg, flow-volume gravity; NICU, neonatal intensive care unit. Values are group means and standard deviations (SD).
or baseline. Two raters, who were both physicians working in the NICU (HG and MB), independently selected breathing segments for computerised analysis using software provided by the manufacturer. The following parameters were recorded for each measurement session and averaged over the selected breathing segment: (i) tidal volume (Vt) in mL, (ii) time to peak tidal expiratory flow as a ratio of total expiratory time (Tptef/Te), and (iii) flow-volume gravity mid-point (FVg) (Fig. 1). FVg is the expiratory flow centroid, that is the dimensionless value given by the position of the centre of gravity of the expiratory flow-volume loop along the volume axis, where the start of expiration is defined as zero and the end as one. A symmetrical, normal curve will give an FVg of 0.5. Airway obstruction will prolong the end-expiratory phase and shift this point to the left and give values below 0.5, as shown in Figure 1. Calculation of the expiratory flow centroid has been found to provide a good description of the tidal expiratory waveform (12) and of changes in the tidal expiratory waveform seen in airflow obstruction (9,11,12).

Comparison of measurements
Repeatability, reproducibility and reliability are important properties when assessing how clinically useful a method is (13–16). Repeatability refers to the variations that occur when test and retest measurements are performed by the same person under identical conditions, and it provides an estimate of the minimum level of agreement between replicated measurements. Reproducibility is the degree of agreement between measurements performed by different examiners or under different conditions, in this case before and after a meal. Reliability is how much of the variations in the measurements are due to the sum of inborn within-subject variation and measurement errors as opposed to true differences in lung function mechanics. If reliability is high, the variability caused by measurement errors and inborn within-subject variations are low compared to the variations caused by subjects being different. Hence, subjects can be well distinguished from each other.

To estimate the repeatability of Nurse A and Nurse B, we compared the measurements of the respective nurses before the meal (A1 before versus A2 before and B1 before versus B2 before) and also after meal for Nurse A (A3 after versus A4 after). To determine the repeatability and reliability for the different categories of infants, comparisons were performed using the results from Nurse A, the most experienced nurse, before the meal. Interexaminer reproducibility was tested by comparing data collected by Nurse A and Nurse B before the meal using the measurements A1 before and A2 before versus B1 before and B2 before. Reproducibility before and after meal was analysed by comparing measurements A1 before and A2 before versus A3 after and A4 after. The results of the computerised analyses after the selection of the two raters were compared. As the breathing segments were obtained by different raters, but from the same measurement sessions, they can be referred to as both interrater repeatability and intrasession repeatability.

Statistical methods
To determine repeatability and reproducibility, we calculated the coefficient of repeatability (CR), which was defined as 1.96 standard deviations (SD) of the mean pairwise differences between measurements. We have reported these findings using CR%, which is CR as a percentage of the pairwise mean. The CR or CR% provides an estimate of the size of the test–retest

Figure 1 Flow-volume gravity mid-point (FVg). FVg is a dimensionless value given by the position of the centre of gravity of the expiratory flow-volume curve along the volume axis, using a linear scale where zero is the volume at start of expiration and one is the volume at end of expiration. FVg will therefore have a value between zero and one. (A) A symmetrical (normal) curve will give an FVg of 0.5. (B) Airway obstruction will shift the curve to the left and produce values for FVg below 0.5. Vg, volume of centre of gravity; Vi, inspiratory volume; Ve, expiratory volume; Vt, tidal volume.
measurement differences, for example a CR% of 10 means that 95% of all the test–retest differences were within ±10% of their pairwise mean. Paired-sample t-tests were applied to assess whether the mean pairwise differences of the various sessions differed significantly from zero. F-tests for the equality of variances were used to detect differences between the CRs. The coefficient of variation (CV) was also calculated for the pairwise measurements. The CV was calculated as the SD of all the measurements divided by their mean and is expressed as a percentage.

To examine reliability, the intraclass correlation coefficient (ICC) was calculated. The ICC is used to assess the consistency, or conformity, of measurements within a group, and it describes how strongly subjects in the same group resemble each other. It is defined as follows:

$$\text{ICC} = \frac{(\text{SD between subjects})^2}{(\text{SD between subjects})^2 + (\text{SD within subjects})^2}$$

The ICC is expressed as a value between zero and one. Reliability was considered good if the ICC values were ≥0.75 and adequate or fair to good if 0.4 < ICC < 0.75 (17). One-way analysis of variance (ANOVA) was used to estimate between-subject and within-subject SDs. Data were analysed with MedCalc version 13.1 (MedCalc Software, Mariakerke, Belgium).

RESULTS
Feasibility and applicability
Having the vests fitted and wearing them did not cause any noticeable discomfort to the infants. There were no changes in respiratory rate, heart rate or oxygen saturation and no subjective changes in facial expressions, breathing patterns or other alterations in behaviour while wearing the vests (data not shown).

Of the 176 recordings, 126 (72%) provided adequate quality for the analyses. The reasons for irregular or variable tidal flow-volume (TFV) loops that could not be analysed were not always obvious, but the system was sensitive to movements from nearby electrical devices, such as spotlights, fluorescent lamps and battery chargers, but these problems were largely solved when we became aware of them. In two participants, measurements were only performed before the meal for practical reasons. A mean of 28 breaths (SD 9) was used to calculate mean values of the different tidal breathing parameters.

Intersession repeatability and reliability
Nurse A obtained successful paired measurements before the meal for 22 participants (A1 before versus A2 before) and Nurse B obtained them for 20 participants (B1 before versus B2 before). After the meal, Nurse A obtained 23 successful paired measurements (A3 after versus A4 after). There were no significant differences between the mean values of any paired measurement sessions for any of the breathing parameters (data not shown).

In terms of CR%, we found relatively large test–retest variability for Vt and Tptef/Te for nurses A and B, which indicates relatively poor repeatability (Table 2). FVg showed good repeatability. Nurse A achieved better repeatability (lower CR%) for all breathing parameters than Nurse B (Table 2), but the difference was only significant for Vt (p < 0.05, results not shown). After meals, the repeatability of Nurse A significantly decreased (CR% increased) for Vt (p < 0.05), but not for the other parameters (Table 2).

CR% tended to be better for the term-born than smaller premature infants, but the differences were not statistically significant (Table 3).

The reliability was good for Vt (ICC > 0.75) for term infants and preterm infants under NICU care and adequate or fair to good (0.4 < ICC < 0.75) for preterm infants at discharge. The opposite pattern was found for Tptef/Te and FVg, where the reliability was good for preterm infants at discharge and adequate or fair to good for term infants and preterm infants receiving NICU care (Table 3).

Reprodicibility for different examiners
There were no significant differences in mean values between the paired measurements performed by nurses A and B, which comprised 42 measurements of A1 before and A2 before versus B1 before and B2 before (Table 4). Interexaminer reproducibility, as expressed by the CR%, was similar to the repeatability of the experienced nurse, Nurse A, before meals (p > 0.05) (data not shown). Again, the CR% showed best reproducibility for FVg and larger variability for the parameters Vt and Tptef/Te.

Reproducibility before and after meal
There were no significant differences in the mean values obtained before and after meals, which comprised 44 measurements of A1 before and A2 before versus A3 after and A4 after (Table 4). However, the reproducibility was poor, especially for Vt and Tptef/Te, as demonstrated by CR% of 81.1 and 78.8, respectively (Table 4).

Interrater and intrasession repeatability
There were no significant differences between the results based on the blinded and independent selections by the two raters from all the 126 recordings (Table 2). The variability expressed by CR% was low for all parameters, suggesting that the raters had little influence on the results.

DISCUSSION
The FloRight system did not disturb the infants or interfere with their respiratory behaviour. The repeatability of the lung function measurements was similar to those reported for traditional infant spirometry using a face mask, which is much more intrusive (5,18–22). The repeatability of the FVg parameter was consistently good and FVg also showed good reliability for preterm infants at discharge. The more
experienced examiner, Nurse A, demonstrated improved repeatability, and the relation to when the infants received their meal appears to have been important.

It may seem confusing that the repeatability was limited despite similar mean values for the various group comparisons, but this apparent contradiction is thoroughly discussed in the statistical literature (13,15,16). Similar mean group values suggested that different examiners, the introduction of a meal or the somewhat extended time frames between measurements did not introduce systematic bias. Moreover, our mean values were basically what we expected when we compared them with those reported by others who used the FloRight system (9,23), except that the preterm-born group still receiving NICU care had smaller weight-corrected tidal volumes in our cohort (Table 1).

There is a scarcity of tidal breathing data in infants under 36 weeks of postmenstrual age, partly due to the complexity of previous measurement methods. Therefore, the influence of factors such as gestational age, weight at birth and at the time of measurement, postnatal age, gender and neonatal lung disease is insufficiently understood and these issues certainly need to be explored further. Some of the inconsistencies regarding tidal volume measurements for electromagnetic inductance plethysmography are somewhat disturbing (9–11). However, even if the method underestimates true tidal volumes, the present data on repeatability and reliability are still helpful and indeed of practical importance.

Test and retest variability and intra-individual variations in lung function measures have commonly been described by the mean CV. In our view, this statistical measure is not a preferable way of expressing repeatability (14), but we used it to compare our results with other studies. The CVs in our study were comparable to those of Lodrup Carlsen et al. (19) and Fuchs et al. (24), which both used a face mask system when examining term-born infants. Olden et al. (9) used the FloRight system on term infants and reported similar repeatability for Vt and FVg, expressed as CR and CR%, as in the present study, but better repeatability for Tptef/Te (CR% 37% versus 51%). We may have improved the variability in Tptef/Te by calculating repeatability over a larger number of breaths (21), but few breaths cannot explain the difference between our results and those of Olden et al. (9), as their results were only based on 10 consecutive breaths.

As our repeatability figures are comparable to those reported for other methods, it is probable that the variability mainly reflects variable breathing patterns of the infants and not the method itself (5,18–22), although, for example, loosening and retightening the vest may have contributed to some of the variability. Newborn infants vary their breathing patterns over time, and this variation is influenced by factors such as gestational age and sleep stage (1). In our study, standardisation of sleep or awake states and shorter time intervals may have improved, but not eliminated, the effect of variations in breathing patterns. Large variations in breathing patterns in infants challenge all attempts to measure infant tidal breathing, and tidal breathing variables

| Table 2 Inter- and intrasession repeatability
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Vt (mL)</td>
<td>0.92 (0.2, 0.3)</td>
<td>79.4* (2.8, 2.3)</td>
</tr>
<tr>
<td>Tptef/Te</td>
<td>4.4 (1.5, 0.4)</td>
<td>13.3</td>
</tr>
<tr>
<td>FVg</td>
<td>0.01 (0.004, 0.03)</td>
<td>13.8</td>
</tr>
</tbody>
</table>

Differences are group means with 95% CIs. CR%, coefficient of repeatability reported as a percentage of the pairwise means. CV, coefficient of variation, expressed as one standard deviation in percentage of the average value of the measurements. Vt, tidal volume; Tptef/Te, time to peak expiratory flow as a ratio of total expiratory time; FVg, flow-volume gravity.

*Significantly higher than CR% for Nurse A before meal (p-value by F-test < 0.05).
should, therefore, be cautiously interpreted irrespective of the measurement technique used. Nevertheless, methods with similar repeatability, such as plethysmographic measures of residual lung volume in children, have proved useful in assessing respiratory disorders (25,26).

We found that the interexaminer reproducibility did not differ from the repeatability of the experienced nurse. However, this does not mean that the experience of the examiner was unimportant. Repeatability was poorer for the six inexperienced nurses, and this showed that experience improved precision. The accuracy of the measurements will depend on the vest fitting properly and on the correct placement of the antenna and the calibration cylinder. Although all of the nurses felt that the system was easy to use, such details may improve with experience.

Our results suggest that the relationship to the meal was important when carrying out lung function testing in infants, as the repeatability after the meal was poorer. This was in contrast to authoritative advice, admittedly based on very limited experience (27). It may be that our finding was particularly relevant for the FloRight method due to abdominal distention and changes in torso topography after a meal. On the other hand, clinical experience suggests that infants breathe more irregularly shortly after meals and it is our opinion that poorer repeatability after a meal was mainly caused by changes in breathing patterns. The similar mean values obtained before and after meals certainly strengthen this assumption.

We found that term-born infants demonstrated a tendency towards better repeatability for all breathing parameters than preterm-born babies. This could be because there were more variable breathing patterns in the preterm-born infants in our cohort (28), but it is also conceivable that a smaller torso may contribute to relatively larger measurement errors. However, the study was vulnerable to type two statistical errors due to the small number of participants. On the other hand, the strengths of our study were the strictly defined patient groups and the meticulous adherence to the study protocol.

The FVg parameter should be investigated more thoroughly, as also pointed out by Olden et al. (9), as it had far better repeatability than Tptef/Te. FVg reflects the shape of the TFV loops, which has been shown to be unchanged during long-term measurements despite breath-to-breath variations (28). By definition, FVg must have a value between zero and one and, in most cases, it is between 0.3 and 0.55, which may limit its capacity to differentiate between health and disease. However, Olden et al. (9) studied prematurely born infants at discharge and showed that FVg differentiated infants with and without bronchopulmonary dysplasia. Their finding corresponds with our finding of better reliability, based on ICC, and larger

| Table 3 | Repeatability and reliability of tidal breathing parameters in the different participant groups (repeated measurements by the same nurse: Nurse A1 versus A2) |
|------------------|---------------------------------|------------------|------------------|
| **Term infants** | **Preterm infants at discharge** | **Preterm infants at under NICU care** |
| CR% | ICC | W-S SD/B-S SD | CV (%) | CR% | ICC | W-S SD/B-S SD | CV (%) | CR% | ICC | W-S SD/B-S SD | CV (%) |
| Vt | 22.2 | 0.98 | 1.4/9.7 | 7.4 | 48.7 | 0.62 | 2.5/3.2 | 17.7 | 47.3 | 0.81 | 1.35/2.71 | 18.7 |
| Tptef/Te | 50.9 | 0.64 | 7.7/10.2 | 22.1 | 61.8 | 0.80 | 8.6/17.3 | 20.8 | 66.0 | 0.59 | 11.7/14.1 | 4.9 |
| FVg | 9.2 | 0.65 | 0.020/0.027 | 4.4 | 14.5 | 0.83 | 0.023/0.051 | 24.6 | 14.9 | 0.55 | 0.027/0.030 | 5.4 |

CR%, coefficient of repeatability reported as a percentage of the pairwise means. ICC, the intraclass correlation coefficient. W-S SD, within-subject standard deviation; B-S SD, between-subject standard deviation; NICU, neonatal intensive care unit.

CV, coefficient of variation, expressed as one standard deviation in percentage of the average value of the measurements.

Vt, tidal volume; Tptef/Te, time to peak expiratory flow as a ratio of total expiratory time; FVg, flow-volume gravity.

Table 4 | Mean values and reproducibility for different nurses (left) and before versus after meal (right) |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nurse A before meal</strong></td>
<td><strong>Nurse B before meal</strong></td>
<td><strong>Difference Nurse A versus Nurse B (95% CI)</strong></td>
<td><strong>CR% reproducibility</strong></td>
</tr>
<tr>
<td>Vt (mL)</td>
<td>13.2 (6.3)</td>
<td>13.3 (6.5)</td>
<td>−0.17 (−1.03, −0.78)</td>
</tr>
<tr>
<td>Tptef/Te</td>
<td>40.9 (12.9)</td>
<td>40.2 (13.4)</td>
<td>0.58 (−3.0, 4.2)</td>
</tr>
<tr>
<td>FVg</td>
<td>0.48 (0.03)</td>
<td>0.47 (0.04)</td>
<td>0.011 (−0.004, 0.025)</td>
</tr>
</tbody>
</table>

Figures are group means with standard deviations (SD) or 95% confidence intervals (95% CI). Nurse A before meal, Nurse B before meal and Nurse A after meal are mean values of two replicate measurements. Differences are measurements made by Nurse A minus Nurse B before meal (right table, n = 42) and Nurse A before meal minus Nurse A after meal (left table, n = 44). CR%, the coefficient of repeatability, reported as a percentage of the pairwise means. Vt, tidal volume; Tptef/Te, time to peak expiratory flow as a ratio of total expiratory time; FVg, flow-volume gravity.
group heterogeneity for FVg, based on larger standard deviations between subjects, in our similar premature group at discharge. This group also had better reliability, based on ICC, for Tptef/Te than the other groups (Table 3). Although all three groups were considered healthy, the premature group at discharge probably had the largest variation in lung function, because of variations in early lung disease and abnormal lung maturation due to prematurity. The finding therefore suggests that the FloRight system can discriminate between patterns of lung function and be useful for exploring lung disease in infancy and early childhood.

The strengths of the FloRight system, as we see it, are its relative simplicity and the fact that it allows continuous and prolonged sampling of data without disturbing the baby and without using a face mask. The method is therefore suitable for lung function testing in large samples of infants (23) with a wide spectrum of respiratory diseases. It can be applied in situations where a face mask is not tolerated, for example in patients requiring continuous positive airway pressure (11,29) or high-flow nasal cannula and if sedation may carry risks, such as in bronchiolitis. The method also appears to be well suited for monitoring the effects of clinical interventions. However, more studies are needed before the system is used widely, particularly in terms of validation against established methods, because there have been few previous validation studies, those that have been carried out have tended to be rather small and they have produced conflicting data.

ACKNOWLEDGEMENTS
We are indebted to Morten Eriksen from VoluSense AS for technical support and for providing Figure 1 in this manuscript. VoluSense AS was not involved in any other aspect of this study.

CONFLICTS OF INTEREST
The authors have no conflict of interests to declare.

FINANCIAL SUPPORT
This study received funding from the Western Norway Regional Health Authority.

References


