



Infant lung function: criteria for selecting tidal flow–volume loops

Karen Eline Stensby Bains^{1,2}, Hrefna Katrín Gudmundsdóttir^{1,2,9}, Martin Färdig^{3,4,9}, Erik Amnö³, Christine M. Jonassen^{5,6}, Björn Nordlund^{1b3,4}, Eva Maria Reh binder^{1b1,7}, Håvard O. Skjerven^{1b1,2}, Corina Silvia Rueegg⁸, Riyas Vettukattil^{1,2} and Karin C. Lødrup Carlsen^{1,2}

¹University of Oslo, Faculty of Medicine, Institute of Clinical Medicine, Oslo, Norway. ²Division of Paediatric and Adolescent Medicine, Oslo University Hospital, Oslo, Norway. ³Astrid Lindgren Children's Hospital, Karolinska University Hospital, Stockholm, Sweden. ⁴Dept of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden. ⁵Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway. ⁶Centre for Laboratory Medicine, Østfold Hospital Trust, Kalnes, Norway. ⁷Dept of Dermatology, Oslo University Hospital, Oslo, Norway. ⁸Oslo Centre for Biostatistics and Epidemiology, Oslo University Hospital, Oslo, Norway. ⁹These authors contributed equally, and share second authorship.

Corresponding author: Karen Eline Stensby Bains (karen.eline.sh@gmail.com)



Shareable abstract (@ERSpublications)

A pre-defined set of inclusion and exclusion criteria provide a reliable and standardised procedure for manual selection of tidal flow–volume loops in infants, which may be useful in clinical and research settings <https://bit.ly/3NYMlp0>

Cite this article as: Bains KES, Gudmundsdóttir HK, Färdig M, *et al.* Infant lung function: criteria for selecting tidal flow–volume loops. *ERJ Open Res* 2022; 8: 00165-2022 [DOI: 10.1183/23120541.00165-2022].

Copyright ©The authors 2022

This version is distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0. For commercial reproduction rights and permissions contact permissions@ersnet.org

Received: 2 April 2022
Accepted: 3 July 2022

Abstract

Background Tidal flow–volume (TFV) loops are commonly recorded in infants during sleep, due to the more regular breathing patterns compared to the awake state. Standardised deselection of loops outside pre-specified ranges are based on periods of regular breathing, while criteria and available software for visual evaluation of TFV loops are lacking. We aimed to determine the reliability of standardised criteria for manual selection of infant TFV loops.

Methods Using a pre-defined set of criteria, three independent raters manually evaluated TFV loops among 57 randomly selected awake healthy 3-month-old infants with available TFV measurements in the Scandinavian Preventing Atopic Dermatitis and ALLergies in children (PreventADALL) study. The TFV loops were sampled using the Eco Medics Exhalyzer D. Criteria for selecting TFV loops included reproducible shape and volume with only one peak in tidal expiratory flow (PTEF), excluding loops with no clear or uneven flow towards PTEF. By intraclass coefficient (ICC), the reliability of agreement between raters was determined for the time to PTEF (t_{PTEF}) to expiratory time (t_E) and other TFV loop parameters.

Results Five infants had unsuccessful tests. Among the remaining 52 infants, the raters selected a median of 25, 26 and 15 loops per test. The ICCs (95% CI) were 0.97 (0.92–0.98) for t_{PTEF}/t_E , 0.99 (0.99–1.00) for respiratory rate, 0.98 (0.97–0.99) for tidal volume per kg and 0.98 (0.97–0.99) for expiratory volume, reflecting excellent agreement in all categories.

Conclusion Manual TFV loop selection using standardised criteria provides a reliable alternative for lung function measures in awake infants with interrupted breathing cycles in a real-life setting.

Introduction

Infant lung function testing has been used to assess lung development and the impact of environmental factors and to detect lung disease. By tracking through childhood and adolescence, lung function in infancy is a major predictor of adult lung function [1–3]. Measures of lung function in awake young children include tidal breath flow–volume loops [4], representing compound measurements of lung function, including size of airways, mechanical characteristics of the lung [5] and respiratory control [6]. During tidal flow–volume (TFV) loop sampling, abnormal patterns of breathing and airway obstruction may be exposed [7]. TFV loop measurements correlate with forced expiratory measurements [8, 9], and lower values of ratio of time to peak tidal expiratory flow (t_{PTEF}) to expiratory time (t_E) in infancy are associated with chronic lung disease, wheeze in infancy and asthma later in life [10–14].



Tidal breathing measures have been obtained in awake and naturally sleeping infants and children, sometimes under sedation [15]. While it is possible to obtain lung function tests from awake preschool children actively participating, historically, it has been challenging to assess lung function in infants without sedation or during natural sleep. Commonly, chloral hydrate has been used as a sedative [16] in inpatient and outpatient facilities [17]; a drug negatively affecting normal ventilation [18], associated with several cases of overdosing, respiratory depression, cardiopulmonary arrest and fatal events [17]. As TFV loops often are more easily obtained during sleep, and breathing cycles are less likely to be interrupted, measurements during sleep has been preferred in infants and young children [8, 19]. However, measurements in the awake state may be advantageous as children are more likely to be awake than sleeping at clinical investigations. Furthermore, lung function in older children is measured in the awake state, and measures obtained in awake compared to sleeping infants may be less influenced by external factors [20]. Associations between t_{P_{TEF}/t_E} and maternal smoking *in utero*, and future asthma are observed in both awake [11, 21] and sleeping states [22, 23]. The clinical value of TFV measures on an individual level is debated [24, 25], partly due to the lack of reference values [26]. Guidelines for TFV measures are established [26], largely based on examination of sleeping or sedated children [19], while sources of variability and criteria for selection of loops are unclear. In the commonly used software, the only option for automatic selection of loops is by pre-defining a threshold for maximum deviation of millilitres from the median tidal volume (V_T).

The American Thoracic Society (ATS)/European Respiratory Society (ERS) guidelines state that automatic breath detection should be accompanied by a visual evaluation of the flow and volume signals; however, there is no clear consensus on criteria for manual inclusion or exclusion of flow–volume loops in a test [26]. There is a need for a validated standard operating procedure, with clear criteria for inclusion and exclusion of TFV loops, for use in clinical as well as research settings, regardless of arousal state. Therefore, the aim of the present study was to determine the reliability of a pre-defined set of criteria for manual selection of TFV loops in infants.

Material and methods

Study subjects

Three independent raters evaluated TFV loop measures in 57 randomly selected infants with available lung function at 3 months of age, antenatally enrolled in the general population-based prospective mother–child birth cohort study Preventing Atopic Dermatitis and ALLergies in children (PreventADALL) [27].

Two raters from Oslo University Hospital and one from Karolinska University Hospital had access to and evaluated TFV measures stored in a secure data server at the University of Oslo. All three raters were medical doctors with clinical experience from general paediatric medicine under training within paediatric pulmonology. Two raters (K.E.S. Bains and H.K. Gudmundsdóttir) both performed infant lung function testing and evaluated the TFV loops, while the third (E. Amnö) participated in the loop evaluation and selection process only.

The PreventADALL study recruited 2697 pregnant women from Norway (Oslo University Hospital and Østfold Hospital Trust) and Sweden (the region of Stockholm) at ~18 weeks of pregnancy from December 2014 to October 2016, and their healthy infants born at or after gestational week 35.0. In the present study, the source population consisted of healthy, awake infants with lung function measurements obtained by study personnel in Oslo.

The PreventADALL study was approved by the regional committee for medical and health research ethics in South-Eastern Norway (2014/518) and Sweden (2014/2242-31-4). The study was registered at ClinicalTrials.gov (NCT02449850). Written informed consent was obtained from the pregnant women at enrolment and from parents at inclusion of the newborn infants.

Procedures

Trained study personnel measured lung function in awake infants at the follow-up examination at 3 months of age. Infants were calm and positioned supine in either a stroller or on a firm pillow on caregiver's arm or lap, with head and neck in midline. The TFV loops were sampled using the Eco Medics Exhalyzer D (Duernten, Switzerland) with ultrasonic flowmeter attached to a carbon dioxide adapter and a dead space reducer (Set 1) with Spirette (Eco Medics) and a tight-fitting face mask with inflated cuff covering nose and mouth to avoid air leaks. The equipment was calibrated daily for atmospheric pressure, temperature and channel, whereas flow calibration was executed between every subject. Analyses with the Spiroware

software version 3.2.1 were in line with international guidelines on infant lung function testing [15, 26]. A test run constituted of consecutive TFV loops to a maximum of 100 loops, as defined as cut-off in the software. Further details on lung function testing are outlined in supplementary material 1.

A set of pre-defined criteria was outlined in a standard operating procedure for manual selection of TFV loops, developed by the raters together with senior researchers in the field of tidal breathing measurements. Details of the standard operating procedure are given in supplementary material 2, with the main criteria illustrated in figures 1 and 2. Briefly, the loops should be reproducible with fairly even shape and similar volumes with only one peak on expiratory flow, while allowing some normal variation as expected in a healthy child. Both consecutive and nonconsecutive breaths were saved when deemed reproducible and with little deviation of volumes from mean V_T during expiration or inspiration. Explicit criteria for exclusion of loops were no clear peak tidal expiratory flow (PTEF) or loops with an aborted or uneven flow towards PTEF at the beginning of the expiratory phase. Each lung function test was eventually rated into one of three pre-defined quality categories: successful, partly successful and not successful. A successful test was defined as a test with good reproducible quality and included preferably ≥ 10 loops. A partly successful test was defined as a test including fewer accepted loops or where reproducibility in loop shape or selected variables was uncertain, showing a greater variance in between the concluded loops. Tests that were not successful were of poor quality, with uncertainty whether the loops represented the infants' normal breathing or included no saved loops.

All three raters independently evaluated, selected and rated all TFV loops sampled in 57 infants and recorded the observations electronically. The infants were randomly selected from a list of all infants attending the 3-month follow-up visit at Oslo University Hospital, using random sampling in SPSS (IBM, Chicago, IL, USA). Each rater independently worked successively through the list to identify infants with lung function measures in the awake state and then evaluated the available loops to classify and qualify each test run. Thereafter, each rater independently deemed the test successful, partly successful or not successful, based upon a general evaluation of the test in relation to the criteria in the standard operating procedure (figure 1 and supplementary material 2). Samples were scrutinised and stored within a safe storage at the Service for Sensitive Data unit at the University of Oslo [27].

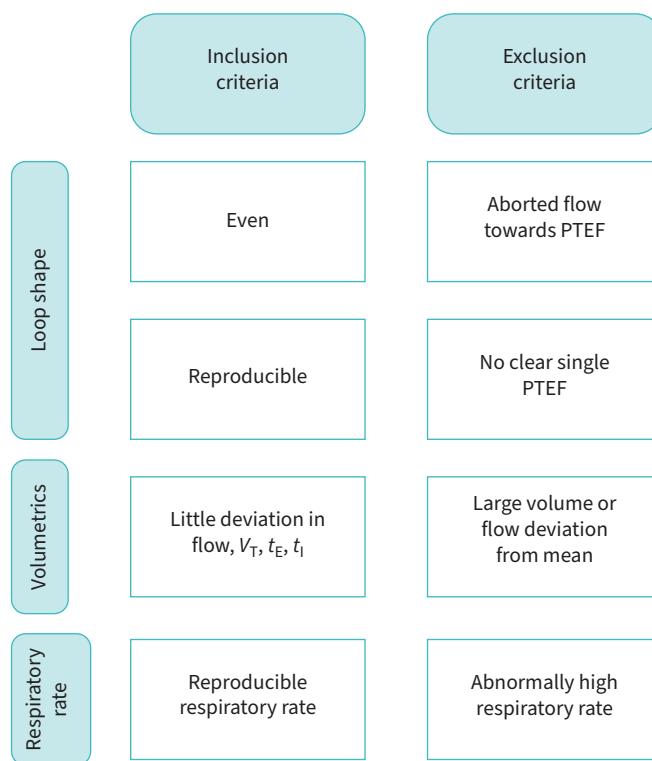


FIGURE 1 Criteria for inclusion and exclusion of loops in a test. V_T : tidal volume; t_E : expiratory time; t_I : inspiratory time; PTEF: peak tidal expiratory flow.

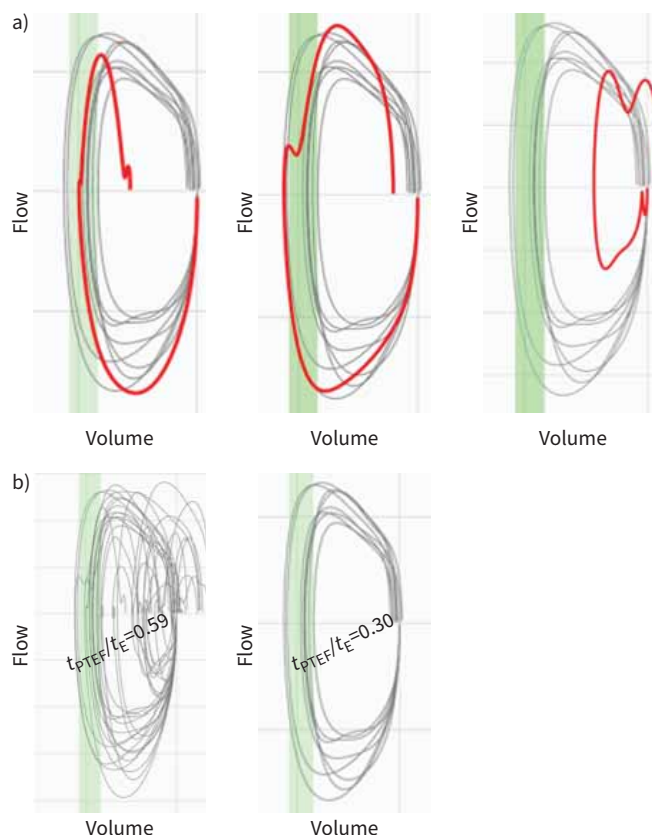


FIGURE 2 Tidal flow–volume loops. a) Examples of excluded loops (red) due to deviating volume (left), notch on expiratory flow (middle) and two expiratory flow peaks (right); b) tidal flow–volume test before (left) and after (right) manual selection of loops. t_{PTEF} : time to peak tidal expiratory flow; t_E : expiratory time.

For comparison of lung function measurement variables between manual and automatic selection of loops, system settings used for automatic selection of loops were set to standard from the manufacturer as described in supplementary material 2. Results of software selected loops were reported electronically without any manual correction.

Definitions and outcomes

The primary outcome was the level of agreement among different raters of t_{PTEF}/t_E , and secondary outcomes were the agreement between the raters of t_{PTEF}/t_E categories <0.20 , $0.20 < 0.25$ and ≥ 0.25 , as well as respiratory rate, tidal volume per kilogram ($V_T \cdot \text{kg}^{-1}$) and expiratory volume (V_E).

Statistical analysis

Descriptive statistics are presented in numbers and proportions for categorical variables and mean or median with standard deviation or minimum and maximum for continuous variables.

For the reliability analysis of the continuous variables, we calculated the intraclass correlation coefficient (ICC) with a two-way random-effects model [28] as we were evaluating a rater-based clinical assessment method where the three raters had similar characteristics, and we planned to generalise our reliability results to other raters. The analysis was based on the single rater type, with the reliability experiment in this article comparing the actual rating of three independent raters. Absolute agreement for the outcomes was assessed by calculating ICC estimates and their 95% confidence intervals using SPSS (version 25). ICC values >0.90 indicate excellent reliability, values between 0.75 and 0.90 good reliability and values between 0.50 and 0.75 indicate moderate reliability [28]. The P0 (the null value of the ICC) was set to 0.6.

Agreement between raters for categorical variables are reported descriptively. We investigated whether the number of loops selected differed between raters using linear regression and calculating robust standard

TABLE 1 Characteristics for the 57 infants with tidal flow–volume measures assessed by three independent raters

Boys	36 (63.2)
Gestational age at birth (weeks)	39.8 (36.6–42.9)
Birthweight (kg)	3.6 (2.8–4.9)
Birth length (cm)	51 (48.0–56.0)
Weight at 3 months (kg)	6.2 (4.57–8.18)
Length at 3 months (cm)	62 (56.5–67.9)
Data are presented as n (%) or mean (range).	

errors to adjust for the cluster “infant”. This analysis was performed in STATA (version 17.0) and included the 156 tests included in the main analysis. To have a statistical power of 80% to detect significant agreement >0.74 , with an α -level of 5% and three raters, the study required TFV loop tests from 53 subjects.

Results

The 57 infants (63% boys) had a mean (range) gestational age of 39.8 (36.6–42.9) weeks and weight at 3 months of age of 6.2 (4.6–8.2) kg (table 1). In five infants no TFV loops were saved after rater assessment by at least one rater and the lung function measurements of these infants were thus not included in the ICC analysis (supplementary table S1). The median number of loops saved by each rater per test among the 52 infants was 25, 26 and 15, respectively (table 2), while the software selected a median of eight loops per test (table 3). In the tests concluded automatic by the software, there were examples of loops with aborted flows and double PTEF. Rater 2 selected on average 2.3 loops more than rater 1 (coefficient 2.3, 95% CI 0.3–4.3; $p<0.001$) and rater 3 selected on average 7.6 loops less than rater 1 (coefficient -7.6 , 95% CI -9.4 – -5.8 ; $p<0.001$).

The mean \pm SD ratios of $t_{\text{PTEF}}/t_{\text{E}}$ were 0.39 ± 0.08 , 0.41 ± 0.08 and 0.39 ± 0.09 for the three raters (table 3), with an ICC of 0.97 (95% CI 0.92–0.98 (figure 3a). The corresponding mean \pm SD ratio of $t_{\text{PTEF}}/t_{\text{E}}$ selected by the software was 0.52 ± 0.22 (table 3).

The ICCs for respiratory rate, $V_{\text{T}}\cdot\text{kg}^{-1}$ and V_{E} were 0.99 (95% CI 0.99–1.0), 0.98 (95% CI 0.97–0.99) and 0.98 (95% CI 0.97–0.99), respectively (figure 3b–d). The ICCs for $t_{\text{PTEF}}/t_{\text{E}}$, respiratory rate, $V_{\text{T}}\cdot\text{kg}^{-1}$ and V_{E} for tests including ≥ 10 loops by each rater ($n=37$) were similar (supplementary table S2).

None of the 52 infants had a $t_{\text{PTEF}}/t_{\text{E}}$ ratio <0.20 , while two infants had a ratio <0.25 as reported by either one or two of the raters. All other infants (96.2%) had a $t_{\text{PTEF}}/t_{\text{E}}$ ratio ≥ 0.25 reported by all raters.

All three raters agreed on the quality category in 41 (72%) out of the 57 infants, with three of these deemed as not successful by all (table 4). Selected TFV measurement parameters for infants where one, two or three raters disagreed on the quality of the test are listed in table 5.

TABLE 2 The number of loops evaluated per rater in the 52 tests that all three raters found appropriate for further analyses, and tests including ≥ 10 loops

	Number of tests	Number of loops
All raters		
Rater 1	52	25 (3–53)
Rater 2	52	26 (3–67)
Rater 3	52	15 (2–52)
Tests including ≥ 10 loops		
Rater 1	45	30 (10–53)
Rater 2	44	29 (11–67)
Rater 3	37	24 (10–52)
Data are presented as n or median (range).		

TABLE 3 Number of loops and ratio of time to peak tidal expiratory flow (t_{PTEF}) to expiratory time (t_E), respiratory rate, tidal volume per kilogram ($V_T \cdot \text{kg}^{-1}$) and expiratory volume (V_E) per rater for infants where data from tests were saved by all three raters, all infants where data were saved by one rater and data selected by the software

	Infants	Number of loops	t_{PTEF}/t_E	Respiratory rate (breaths·min ⁻¹)	$V_T \cdot \text{kg}^{-1}$	V_E (mL)
All raters						
Rater 1	52	25 (3–53)	0.39±0.08	67±13.4	6.4±2.1	39.1±12.2
Rater 2	52	26 (3–67)	0.41±0.08	66±13.1	6.7±2.1	40.2±12.2
Rater 3	52	15 (2–52)	0.39±0.09	65±13.6	6.9±2.1	41.5±12.0
Data saved by one rater						
Rater 1	54	23 (3–53)	0.39±0.08	67±13.3	6.3±2.2	38.4±12.5
Rater 2	56	25 (1–67)	0.43±0.11	66±13.0	6.4±2.2	38.9±12.9
Rater 3	52	15 (2–52)	0.39±0.09	65±13.6	6.9±2.1	41.5±12.0
Spioware	56	8 (1–44)	0.52±0.22	64±14.4	5.4±2.0	29.3±10.2

Data are presented as n, median (range) or mean±SD.

Discussion

The reliability of a set of pre-defined criteria for manual selection of TFV parameters among healthy 3-month-old infants was excellent, with an ICC of 0.97 (95% CI 0.92–0.98) for the t_{PTEF}/t_E ratio between three independent raters. Likewise, the ICC was higher than 0.90 for respiratory rate, $V_T \cdot \text{kg}^{-1}$ and V_E .

To the best of our knowledge, this is the first study to validate a set of pre-defined criteria for individually evaluating TFV loops in awake infants, showing excellent agreement between three independent raters on the t_{PTEF}/t_E ratio, despite varying number of loops approved by each rater. In 25 sleeping infants in the first week of life, YUKSEL *et al.* [25] reported good interobserver repeatability between two observers of t_{PTEF}/t_E measured in a whole-body plethysmograph, using the method of Bland and Altman.

While none of the infants had a $t_{PTEF}/t_E < 0.20$, two infants were rated to have a t_{PTEF}/t_E between 0.20 and 0.25 by one or two raters, while all raters agreed on a t_{PTEF}/t_E ratio of ≥ 0.25 in 52 (96.3%) out of the 54 infants. A cut-off value of t_{PTEF}/t_E ratio < 0.20 is associated with later bronchial obstruction [11, 29], whereas ratio values of ≥ 0.25 have been regarded normal [8, 14, 25, 29, 30]. While low t_{PTEF}/t_E appears clinically relevant, the PreventADALL study is based on a normal population and we assumed, as observed, that the majority of included infants would have lung function values in the normal range. Therefore, we categorised the t_{PTEF}/t_E ratio into three categories of low (< 0.20), marginal ($0.20 < 0.25$) and normal (≥ 0.25) values for comparison, in addition to the exact values included in the ICC analyses.

The ICC for t_{PTEF}/t_E was consistent across a varying number of tidal breathing loops, with the three raters approving a median of 15, 25 and 26 loops, where on average rater 2 saved more and rater 3 saved fewer loops per test than rater 1. The automatic selection of loops by the software in general resulted in fewer loops, with a smaller range, higher mean ratio of the t_{PTEF}/t_E , lower mean $V_T \cdot \text{kg}^{-1}$ and V_E and similar mean respiratory rate as compared to the three raters. The large discrepancy indicates that faulty loops such as double peaks and irregular shaped loops were not deselected by the automatic process. We therefore suggest that manually selected loops are more likely to be representative and of higher quality than are loops selected automatically. STOCKS *et al.* [19] suggested in 1994 that for infants aged > 6 weeks, 10 breath loops might be adequate, whereas in younger infants a t_{PTEF}/t_E based on the mean of 15–20 loops would be a closer estimate to their true value due to decreased within-subject variability with increased age. However, a study has documented tidal flow–volume loop indices based on only four loops selected from a preview of eight loops, due to data storage capacity at the time [31]. The ATS/ERS guidelines on pulmonary function testing for preschool children suggest that a reliable t_{PTEF}/t_E should be based on ≥ 10 loops [15] and these criteria are widely used for TFV measurements in infants as well. However, based on our results, including varying number of included loops by the raters, as shown in table 2, we suggest that tests may be valid even with < 10 loops.

We included tests deemed successful, partly successful and not successful and found a consistent ICC of excellent agreement in both t_{PTEF}/t_E , as well as respiratory rate, $V_T \cdot \text{kg}^{-1}$ and V_E . Defining the quality of a single test was based on the visual shape and reproducibility of the loops, after manual removal of loops with poor technical quality or without a well-defined PTEF. There are no clear-cut criteria for evaluation of the quality of the tests; however, the criteria are outlined in the standard operating procedure for lung

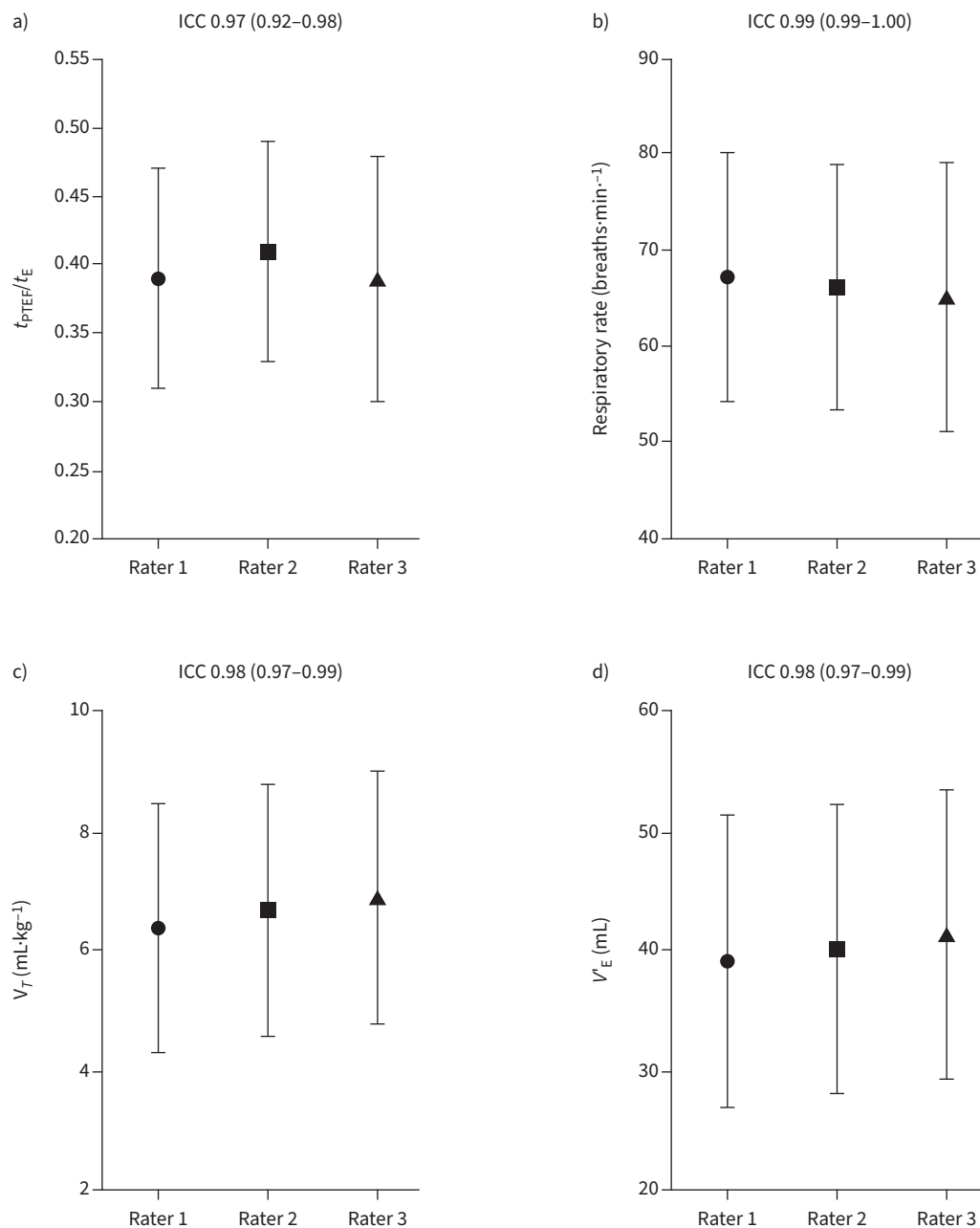


FIGURE 3 Rater agreement with the mean intraclass correlation (ICC) (95% CI) for a) ratio of time to peak tidal expiratory flow (t_{PTEF}/t_E), b) respiratory rate, c) tidal volume (V_T) and d) expiratory volume (V_E). Data are presented as individual mean±SD.

TABLE 4 Conclusion by raters on quality of tests from all 57 infants

	Successful	Partly successful	Not successful	Total number of infants
3 raters agreed	37	1	3	41
2 raters agreed	5	8	0	13
No agreement	4 [#]	8 [#]	10 [#]	3

[#]: concluded by one rater.

TABLE 5 An overview of all infants with rater disagreement on the quality of the test

Infant	Quality of test	Loops before selection	Loops after selection	t_{PTEF}/t_E	$V_T \cdot \text{kg}^{-1}$ (mL)
1					
Rater 1	Successful	60	30	0.44	4.6
Rater 2	Successful		22	0.49	5.4
Rater 3	Partly successful		13	0.47	5.6
6					
Rater 1	Partly successful	44	5	0.44	6.6
Rater 2	Partly successful		9	0.55	6.9
Rater 3	Not successful		3	0.45	8.3
8					
Rater 1	Partly successful	32	8	0.4	2.4
Rater 2	Partly successful		8	0.37	2.4
Rater 3	Not successful		0		
9					
Rater 1	Partly successful	38	3	0.46	7.4
Rater 2	Partly successful		5	0.53	6.5
Rater 3	Not successful		2	0.51	7.7
13					
Rater 1	Partly successful	60	6	0.38	11.6
Rater 2	Successful		15	0.44	12.1
Rater 3	Partly successful		7	0.39	11.8
15					
Rater 1	Successful	60	19	0.54	7.3
Rater 2	Successful		26	0.57	7.4
Rater 3	Partly successful		8	0.52	7.4
22					
Rater 1	Partly successful	60	5	0.31	4.9
Rater 2	Partly successful		7	0.35	5.1
Rater 3	Not successful		4	0.25	5.4
24					
Rater 1	Successful	60	12	0.41	6.7
Rater 2	Successful		20	0.46	8.4
Rater 3	Partly successful		5	0.37	7.4
35					
Rater 1	Successful	60	18	0.45	4.9
Rater 2	Partly successful		6	0.45	6.1
Rater 3	Not successful		3	0.51	7.1
39					
Rater 1	Partly successful	14	3	0.44	3.3
Rater 2	Partly successful		4	0.47	3.8
Rater 3	Not successful		0		
40					
Rater 1	Successful	38	11	0.39	5.9
Rater 2	Partly successful		11	0.47	6.4
Rater 3	Not successful		3	0.47	6
41					
Rater 1	Successful	40	11	0.31	5.5
Rater 2	Successful		13	0.36	5.2
Rater 3	Partly successful		7	0.29	6.1
47					
Rater 1	Partly successful	42	5	0.26	4.4
Rater 2	Partly successful		4	0.32	5.3
Rater 3	Not successful		3	0.26	4.6
48					
Rater 1	Successful	50	12	0.6	6.9
Rater 2	Successful		11	0.59	7.4
Rater 3	Partly successful		5	0.6	7.5
51					
Rater 1	Successful	56	10	0.38	2.9
Rater 2	Partly successful		7	0.51	4.6
Rater 3	Not successful		4	0.39	5.7

Continued

TABLE 5 Continued

Infant	Quality of test	Loops before selection	Loops after selection	t_{PTEF}/t_E	$V_T \cdot \text{kg}^{-1}$ (mL)
53					
Rater 1	Partly successful	31	3	0.34	1.6
Rater 2	Partly successful		3	0.34	1.6
Rater 3	Not successful		2	0.32	1.7
t_{PTEF} : time to peak tidal expiratory flow; t_E : expiratory time; $V_T \cdot \text{kg}^{-1}$: tidal volume per kilogram.					

function analysis in the study provided in online supplement 2. This can explain why the raters deemed tests as being of different quality categories and saved different numbers of loops, where rater 3 on a general basis saved fewer loops from the tests and in general deemed more tests partly successful compared to the other two raters. Despite these discrepancies, the ICC is high for the continuous variables.

The criteria for manual selection of TFV curves in a real-life setting provided sufficiently robust criteria for excellent agreement on the tests, supporting the usefulness of the criteria.

We are not aware of other studies that compare lung function variables from awake TFV measurements manually evaluated by several independent raters.

Strengths and limitations

The TFV tests were performed on healthy awake infants with characteristics reflecting a normal population under standardised circumstances by trained personnel [5, 24]. The number of infants was pre-defined by power calculations to include sufficient numbers of tests. The study requirement of 53 infants to ensure a statistical power of 80% to detect significant agreement >0.74 was not met; however, with 52 infants included in calculations resulting in an ICC >0.90 for all variables, it is unlikely that including one more infant would affect the outcome.

There was little variance in the pre-defined categories of ratios of <0.20 , $0.20 < 0.25$ and ≥ 0.25 defined to distinguish an assumed healthy infant from an infant with reduced lung function. The category based on t_{PTEF}/t_E $0.20 < 0.25$ is somewhat arbitrary, but was pre-defined as being in the lower range of presumably normal TFV loops. The high ICC, reflecting excellent agreement between different raters, was evident for the continuous variables in all outcomes.

The present study provides a further step to standardise TFV measures in epidemiologic studies and clinical practice, in line with the need for further insight into lung function measurement techniques, allowing repeated measurements in awake young children [4]. It remains unclear if selecting loops by this method will be useful in the clinical practice, and further studies should be conducted to validate use for long-term care of patients.

Conclusion

Using a set of pre-defined selection criteria, manual selection of TFV loops from healthy awake 3-month-old infants resulted in excellent agreement of TFV parameters between three independent raters. Our study provides a feasible and valid tool for selecting TFV measures in infants, which may particularly be useful in the absence of long sequences of regular breathing, such as in daily clinical practice.

Provenance: Submitted article, peer reviewed.

Acknowledgements: We sincerely thank all families participating in the PreventADALL (Preventing Atopic Dermatitis and Allergies in Children) study as well as study personnel who have contributed to recruiting participants, performing lung function measurements and other relevant examinations, and in generally managing the study: Mari Kjendsli, Ingvild Essén, Malén Gudbrandsgard (Division of Paediatric and Adolescent Medicine, Oslo University Hospital, Oslo, Norway); Live S. Nordhagen (Division of Paediatric and Adolescent Medicine, Oslo University Hospital and VID Specialized University, Oslo, Norway); Sandra Götberg, Kajsa Sedergrén, Natasha Sedergrén, Caroline-Aleksis Olsson Mägi, Sandra G. Tedner and Ellen Tegnerud (Astrid Lindgren Children's Hospital, Karolinska University Hospital and Dept of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden).

The study was performed within ORACLE (the Oslo Research Group of Asthma and Allergy in Childhood; the Lung and Environment).

This study is registered at www.clinicaltrials.gov with identifier number NCT02449850.

Author contributions: All authors have contributed substantially to the design and/or data collection and/or clinical follow-up of the PreventADALL study, revised the work critically for important intellectual content and approved the final version before submission.

Conflict of interest: E.M. Reh binder has received honoraria for lectures from Sanofi-Genzyme, Novartis, Leo-Pharma, Perrigo and The Norwegian Asthma and Allergy Association. The other authors have nothing to disclose.

Support statement: The PreventADALL study has been funded by The Regional Health Board South East, The Norwegian Research Council, Oslo University Hospital, the University of Oslo, Health and Rehabilitation Norway, the Foundation for Healthcare and Allergy Research in Sweden (*Vårdalstiftelsen*), the Swedish Asthma and Allergy Association's Research Foundation, Swedish Research Council – the Initiative for Clinical Therapy Research, The Swedish Heart–Lung Foundation, SFO-V Karolinska Institutet, Østfold Hospital Trust, the European Union (MedALL project), by unrestricted grants from the Norwegian Association of Asthma and Allergy, the Kloster FOUNDATION, Thermo-Fisher Uppsala, Sweden, by supplying allergen reagents, the Norwegian Society of Dermatology and Venerology, Arne Ingel Legat. Funding information for this article has been deposited with the Crossref Funder Registry.

References

- 1 Stern DA, Morgan WJ, Wright AL, *et al.* Poor airway function in early infancy and lung function by age 22 years: a non-selective longitudinal cohort study. *Lancet* 2007; 370: 758–764.
- 2 Martinez FD. Early-life origins of chronic obstructive pulmonary disease. *N Engl J Med* 2016; 375: 871–878.
- 3 Stocks J, Hislop A, Sonnappa S. Early lung development: lifelong effect on respiratory health and disease. *Lancet Respir Med* 2013; 1: 728–742.
- 4 Moeller A, Carlsen KH, Sly PD, *et al.* Monitoring asthma in childhood: lung function, bronchial responsiveness and inflammation. *Eur Respir Rev* 2015; 24: 204–215.
- 5 van der Ent CK, Brackel HJ, van der Laag J, *et al.* Tidal breathing analysis as a measure of airway obstruction in children three years of age and older. *Am J Respir Crit Care Med* 1996; 153: 1253–1258.
- 6 Seddon PC, Davis GM, Coates AL. Do tidal expiratory flow patterns reflect lung mechanics in infants? *Am J Respir Crit Care Med* 1996; 153: 1248–1252.
- 7 Pijnenburg MW, Baraldi E, Brand PL, *et al.* Monitoring asthma in children. *Eur Respir J* 2015; 45: 906–925.
- 8 Hevroni A, Goldman A, Blank-Brachfeld M, *et al.* Use of tidal breathing curves for evaluating expiratory airway obstruction in infants. *J Asthma* 2018; 55: 1331–1337.
- 9 Morris MJ, Lane DJ. Tidal expiratory flow patterns in airflow obstruction. *Thorax* 1981; 36: 135–142.
- 10 Schmalisch G, Wilitzki S, Wauer RR. Differences in tidal breathing between infants with chronic lung diseases and healthy controls. *BMC Pediatr* 2005; 5: 36.
- 11 Håland G, Carlsen KC, Sandvik L, *et al.* Reduced lung function at birth and the risk of asthma at 10 years of age. *N Engl J Med* 2006; 355: 1682–1689.
- 12 Lødrup Carlsen KC, Carlsen KH, Nafstad P, *et al.* Perinatal risk factors for recurrent wheeze in early life. *Pediatr Allergy Immunol* 1999; 10: 89–95.
- 13 Hoo AF, Henschen M, Dezateux C, *et al.* Respiratory function among preterm infants whose mothers smoked during pregnancy. *Am J Respir Crit Care Med* 1998; 158: 700–705.
- 14 Young S, Arnott J, Le Souef PN, *et al.* Flow limitation during tidal expiration in symptom-free infants and the subsequent development of asthma. *J Pediatr* 1994; 124: 681–688.
- 15 Beydon N, Davis SD, Lombardi E, *et al.* An official American Thoracic Society/European Respiratory Society statement: pulmonary function testing in preschool children. *Am J Respir Crit Care Med* 2007; 175: 1304–1345.
- 16 Lesnick BL, Davis SD. Infant pulmonary function testing: overview of technology and practical considerations – new current procedural terminology codes effective 2010. *Chest* 2011; 139: 1197–1202.
- 17 Paparella SF. Chloral hydrate: safety risks still worth mentioning. *J Emerg Nurs* 2018; 44: 81–83.
- 18 Aldrete JA, Itkin IH. Effects of chloral hydrate on the respiration of nonasthmatic and asthmatic patients. *J Allergy* 1969; 43: 342–348.
- 19 Stocks J, Dezateux CA, Jackson EA, *et al.* Analysis of tidal breathing parameters in infancy: how variable is TPTEF:TE? *Am J Respir Crit Care Med* 1994; 150: 1347–1354.
- 20 Lødrup Carlsen KC, Magnus P, Carlsen K-H. Lung function by tidal breathing in awake healthy newborn infants. *Eur Respir J* 1994; 7: 1600–1668.
- 21 Lødrup Carlsen KC, Jaakkola JJ, Nafstad P, *et al.* *In utero* exposure to cigarette smoking influences lung function at birth. *Eur Respir J* 1997; 10: 1774–1779.

- 22 Stick SM, Burton PR, Gurrin L, *et al.* Effects of maternal smoking during pregnancy and a family history of asthma on respiratory function in newborn infants. *Lancet* 1996; 348: 1060–1064.
- 23 Turner S, Fielding S, Mullane D, *et al.* A longitudinal study of lung function from 1 month to 18 years of age. *Thorax* 2014; 69: 1015–1020.
- 24 Carlsen KC. Tidal breathing measurements. *In:* Hammer J, Eber E, eds. Paediatric Pulmonary Function Testing. Basel, Karger, 2005; pp. 10–19.
- 25 Yuksel B, Greenough A, Giffin F, *et al.* Tidal breathing parameters in the first week of life and subsequent cough and wheeze. *Thorax* 1996; 51: 815–818.
- 26 Bates JH, Schmalisch G, Filbrun D, *et al.* Tidal breath analysis for infant pulmonary function testing. ERS/ATS Task Force on Standards for Infant Respiratory Function Testing. European Respiratory Society/American Thoracic Society. *Eur Respir J* 2000; 16: 1180–1192.
- 27 Lødrup Carlsen KC, Reh binder EM, Skjerven HO, *et al.* Preventing Atopic Dermatitis and ALLergies in Children – the PreventADALL study. *Allergy* 2018; 73: 2063–2070.
- 28 Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016; 15: 155–163.
- 29 Lodrup Carlsen KC, Carlsen KH, Nafstad P, *et al.* Perinatal risk factors for recurrent wheeze in early life. *Pediatr Allergy Immunol* 1999; 10: 89–95.
- 30 Martinez FD, Morgan WJ, Wright AL, *et al.* Diminished lung function as a predisposing factor for wheezing respiratory illness in infants. *N Engl J Med* 1988; 319: 1112–1117.
- 31 Lødrup Carlsen KC, Stenzler A, Carlsen KH. Determinants of tidal flow volume loop indices in neonates and children with and without asthma. *Pediatr Pulmonol* 1997; 24: 391–396.