Differences in lung clearance index and functional residual capacity between two commercial multiple-breath nitrogen washout devices in healthy children and adults

Zwitserloot, Annelies M; van den Born, Evelyne J; Raaijmakers, Lena H A; Stalman, Wouter E; van Smaalen, Marjanne; van den Berge, Maarten; Gappa, Monika; Koppelman, Gerard H; Willemse, Brigitte W M

Published in:
ERJ Open Research

DOI:
10.1183/23120541.00247-2019

Publication date:
2020

ORIGINAL ARTICLE
LUNG FUNCTION

Differences in lung clearance index and functional residual capacity between two commercial multiple-breath nitrogen washout devices in healthy children and adults

Annelies M. Zwitserloot[1,2], Evelyne J. van den Born[1], Lena H.A. Raaijmakers[3], Wouter E. Stalman[1], Marjanne van Smaalen[1], Maarten van den Berge[2,4], Monika Gappa[5], Gerard H. Koppelman[1,2] and Brigitte W.M. Willemse[1,2]

Affiliations: [1]University of Groningen, University Medical Center Groningen, Beatrix Children's Hospital, Dept of Paediatric Pulmonology and Paediatric Allergy, Groningen, The Netherlands. [2]University of Groningen, University Medical Center Groningen, Groningen Research Institute for Asthma and COPD, Groningen, The Netherlands. [3]Radboud University Medical Center, Radboud Institute for Health Sciences, Dept of Primary and Community Care, Nijmegen, The Netherlands. [4]University of Groningen, University Medical Center Groningen, Dept of Pulmonary Diseases, Groningen, The Netherlands. [5]Evangelisches Krankenhaus Düsseldorf, Children's Hospital, Düsseldorf, Germany.

Correspondence: Annelies M. Zwitserloot, Dept of Pediatric Pulmonology and Pediatric Allergy, Beatrix Children's Hospital, University Medical Center Groningen, PO Box 30.001, CA43, Groningen, The Netherlands. E-mail: a.m.zwitserloot@umcg.nl

ABSTRACT Multiple-breath nitrogen washout (MBNW) and its clinical parameter lung clearance index (LCI) are gaining increasing attention for the assessment of small airway function. Measurement of LCI relies on accurate assessment of functional residual capacity (FRC). The EasyOne Pro LAB (ndd) and Exhalyzer D (EM) are two commercially available MBNW devices. The aim of the study was to compare these two devices *in vitro* and *in vivo* in healthy subjects with regard to FRC, LCI and secondary outcome parameters and to relate FRC_{MBNW} to FRC measured by body plethysmography (pleth) and helium dilution technique. MBNW measurements were performed using a lung model (FRC between 500 and 4000 mL) *in vitro* and in 38 subjects aged 6–65 years followed by helium dilution and pleth *in vivo* using fixed and relaxed breathing techniques. *In vitro* accuracy within 5% of lung model FRC was 67.3% for ndd, FRC was >5% higher for EM in all tests. *In vivo*, FRC_{pleth} ranged from 1.2 to 5.6 L. Mean differences (limits of agreement) between FRC_{pleth} and FRC_{MBNW} were −7.0%, (−23.2 to 9.2%) and 5.7% (−11.2 to 22.6%) using ndd and EM, respectively. FRC_{ndd} was consistently lower than FRC_{EM} (−11.8% (−25.6 to 2%)). LCI was comparable between the two devices (−1.3% (−21.9 to 19.3%)). There was a difference of >10 % in LCI in 12 of 38 subjects. Using the most recent software updates, both devices show relevant deviations in FRC measurement both *in vitro* and *in vivo* and individual differences in LCI in a significant proportion of subjects. The devices are therefore not interchangeable.

@ERSpublications
MBNW measurements with the Exhalyzer D and EasyOne Pro LAB cannot be used interchangeably for FRC and LCI measurements. FRC measured on both devices showed deviations from *in vitro* and *in vivo* measurements. https://bit.ly/2xyyUuJ

Cite this article as: Zwitserloot AM, van den Born EJ, Raaijmakers LHA, *et al*. Differences in lung clearance index and functional residual capacity between two commercial multiple-breath nitrogen washout devices in healthy children and adults. *ERJ Open Res* 2020; 6: 00247-2019 [https://doi.org/10.1183/23120541.00247-2019].

This article has supplementary material available from openres.ersjournals.com

Received: 12 Sept 2019 | Accepted after revision: 7 April 2020

Introduction

Lung diseases such as cystic fibrosis (CF) and asthma may have their onset in the small airways [1–3]. Therefore, considerable interest exists in measuring small airways dysfunction, using the multiple-breath nitrogen washout (MBNW) technique. The lung clearance index (LCI) is its most straightforward outcome parameter. The LCI has increasingly gained attention as a sensitive marker of small airways dysfunction and ventilation inhomogeneity in patients with lung disease, especially in CF [4–10]. The reliability of the LCI depends directly on correct measurement of functional residual capacity (FRC), because LCI is calculated as the number of lung volume turnovers (cumulative exhaled volume/FRC) required to lower the tracer gas concentration to 1/40 of the initial concentration [11].

Currently, two commercially available devices are used for MBNW using N_2 as an inert tracer gas: the EasyOne Pro LAB ("ndd"; ndd Medical Technologies, Zurich, Switzerland) and the Exhalyzer D ("EM"; Eco Medics AG, Duernten, Switzerland). In previous studies, these devices have been compared *in vitro* using a lung model and to other techniques to measure FRC such as the helium dilution technique (FRC_{He}) and body plethysmography (FRC_{pleth}) [12–19]. The results from these studies suggest that the ndd measures significantly lower FRC values compared to body plethysmography, whereas EM measures similar to significantly higher FRC values compared to body plethysmography. Limitations of these studies were: use of only one of the two devices [12, 13, 15, 16], inclusion of a limited range of FRC by measuring either children or adults and/or use of an old software version of the ndd [14, 16–20]. Therefore, in this study we directly compared both devices with each other with the most recent software version of ndd and included both healthy children and adults.

The effect of breathing pattern on LCI and FRC is not well examined. Relaxed steady tidal breathing has historically been used for measuring the LCI. To derive other indices, such as parameters derived from (normalised) slope III (SnIII) analysis to distinguish between global ventilation inhomogeneity and specific location of ventilation inhomogeneity, a fixed tidal volume is advocated in the European Respiratory Society (ERS)/American Thoracic Society (ATS) statement [11, 21–23]. A tidal volume (V_T) between 950 and 1400 mL·kg^{-1} for adults and between 10 and 15 mL·kg^{-1} for children is mandatory as a compromise between maintaining physiological breathing and having sufficient phase III to gain SnIII data [11]. It is not exactly known how these different breathing patterns influence the outcome of FRC and LCI. Previous studies that investigated the influence of the breathing pattern on MBNW outcome parameters did not use the volumes recommended by the ATS/ERS statement, and none investigated both adults and children [24, 25].

The aim of the present study was to compare two currently available commercial MBNW devices (ndd and EM), using FRC and LCI as the main outcome parameters in healthy volunteers aged 6–65 years. Additionally, we evaluated the influence of fixed and relaxed tidal breathing patterns on FRC and LCI results.

Methods

Detailed description of randomisation, recruitment and inclusion criteria, study protocol, software versions, power calculation and statistical analysis are provided in the online supplementary material.

Study design

This is a cross-sectional randomised study of *in vitro* and *in vivo* MBNW measurements.

In vitro *MBNW testing*

A lung model setup (Soloplex, Tidaholm, Sweden) based on the model used by Singer *et al.* [13] was used to conduct *in vitro* experiments. With one modification: continuous CO_2 was not added to the model. To simulate different breathing patterns, a positive airway pressure ventilator with volume-controlled mode (Breas PV 501, Breas Medical AB, Mölnlycke, Sweden) was connected to the lung model. All measurements were performed by two operators (AZ and LR). To compare lung volume ranges and breathing patterns observed in subjects aged 6–65 years the target FRC of the lung model was set between 500 and 4000 mL, with a respiratory rate between 10 and 30 per min and V_T between 300 and 1000 mL. All tests were performed in triplicate, in randomised order of target FRC on both devices.

In vivo *MBNW testing*

Setting and participants

Healthy volunteers aged between 6 and 65 years were enrolled between April 2016 and April 2017 to conduct MBNW at the lung function department of the Beatrix Children's' Hospital, University Medical Centre Groningen (Groningen, the Netherlands). The study was approved by the local ethics committee (METc 2015.417). Written informed (parental) consent from all participants was obtained.

Protocol

The MBNW devices used were the ndd and EM devices. N_2 was used as a tracer gas. Randomisation took place for the order of the MBNW device used (ndd or EM) and the pattern of breathing. Breathing pattern (fixed or relaxed) was defined in agreement with the ATS/ERS consensus statement (figure 1) [11].

After the MBNW measurements lung function tests were performed in a fixed order: helium dilution technique (MasterScreen PFT System, CareFusion, USA), spirometry (MasterScreen pneumo-spirometer, CareFusion, USA) and pleth (MasterScreen Body Plethysmography, CareFusion, USA) (figure 1). All tests were performed according to the ATS/ERS guidelines and consensus statement [11, 26, 27].

Acceptable limits and statistical analysis

Minimal within-test repeatability of FRC measured by He, pleth and MBNW is defined as 10%, 5% and 10%, respectively [11, 27]. FRC measured using pleth includes the physiological amount of trapped gas, therefore FRC may be up to 10% higher than FRC measured by gas washout [28, 29]. Taking these two facts into account, we defined acceptable differences between FRC_{MBNW} and FRC_{pleth} as within the limits of −10% to 5%, and FRC_{MBNW} and FRC_{He} as within the limits of −10 and 10%.

Statistical analysis was performed using SPSS 23.0 for windows (IBM SPSS Statistics, version 23.0, Armonk, NY, USA). Statistical significance was set at $p<0.05$.

Accuracy of *in vitro* data was assessed according to the consensus statement: at least 95% of measured FRC should be within 5% of the lung model FRC [11]. The agreement between the devices, and different

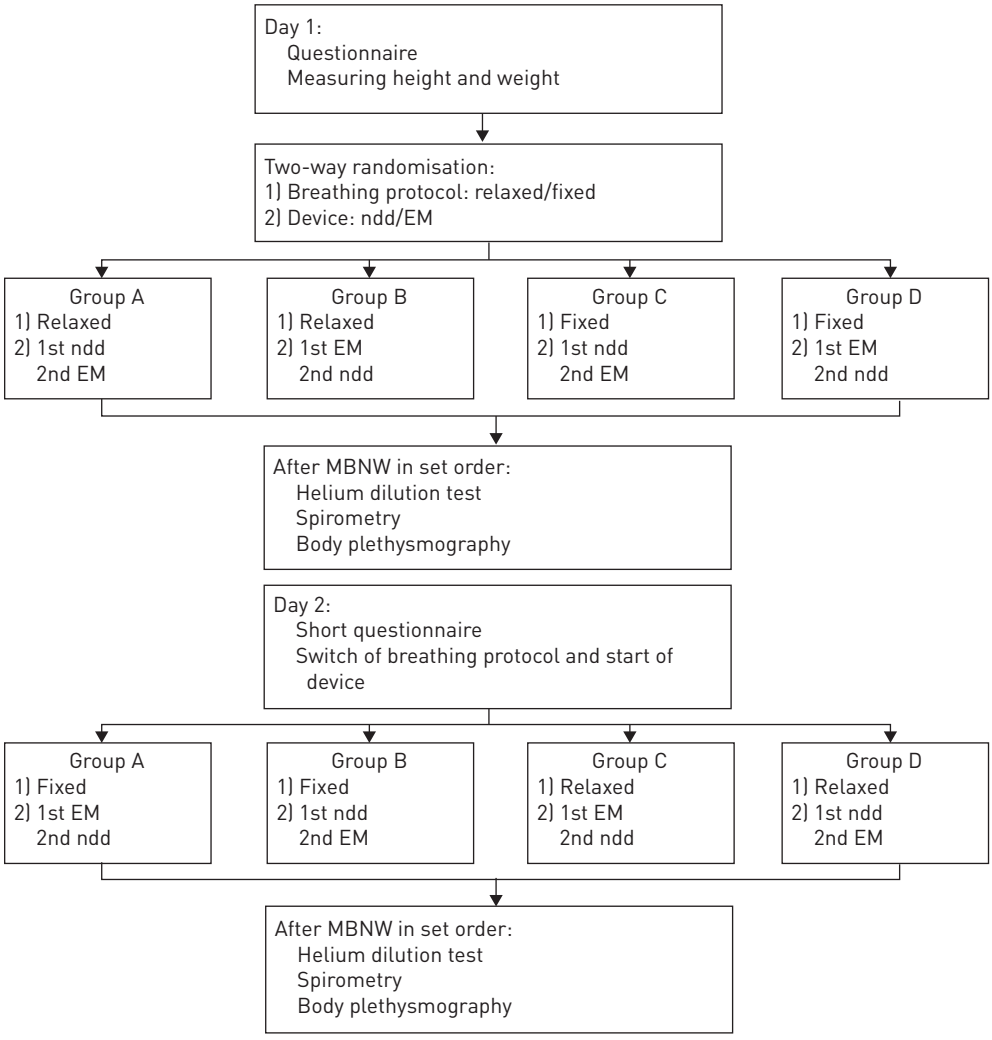

FIGURE 1 Flow diagram study design. Relaxed: relaxed tidal breathing; Fixed: fixed tidal breathing (adults 950–1400 mL, children 10–15 $mL \cdot kg^{-1}$); EM: Exhalyzer D; ndd: EasyOne Pro LAB; MBNW: multiple-breath nitrogen washout.

breathing manoeuvres was assessed using the method by Bland and Altman [30]. According to the consensus statement on MBNW, relaxed tidal breathing was used to compare the FRC_{ndd} and FRC_{EM} with the other lung function tests that measure FRC [11].

Results

In vitro *MBNW testing*

Thirty-four tests, each in triplicate, were performed per MBNW device. Mean difference between FRC measured using ndd compared to $FRC_{lungmodel}$ was 4% (p=0.78, limits of agreement of −8.5 to 12.7%). The accuracy rate of 5% between FRC_{ndd} and $FRC_{lungmodel}$ was reached in 67.6%. For FRC_{EM}, mean difference was 14.3% (p≤0.01) with limits of agreement of −2.2 to 30.8% (figure 2, table 1). All $FRC_{lungmodel}$ measurements with EM were >5% higher and did not reach the defined accuracy rate. The *in vitro* coefficient of variation of FRC was 0.7% using the ndd and 1.2% using the EM.

In vivo *MBNW testing*

A total of 44 healthy volunteers were enrolled between April 2016 and May 2017; 23 children (6–17 years) and 21 adults (18–65 years). Five children, aged 6 to 10 years, were excluded: four were not able to perform acceptable MBNW tests and one showed signs of obstruction on spirometry and had a bronchodilator response of 18%, without respiratory complaints. One adult was not able to perform the

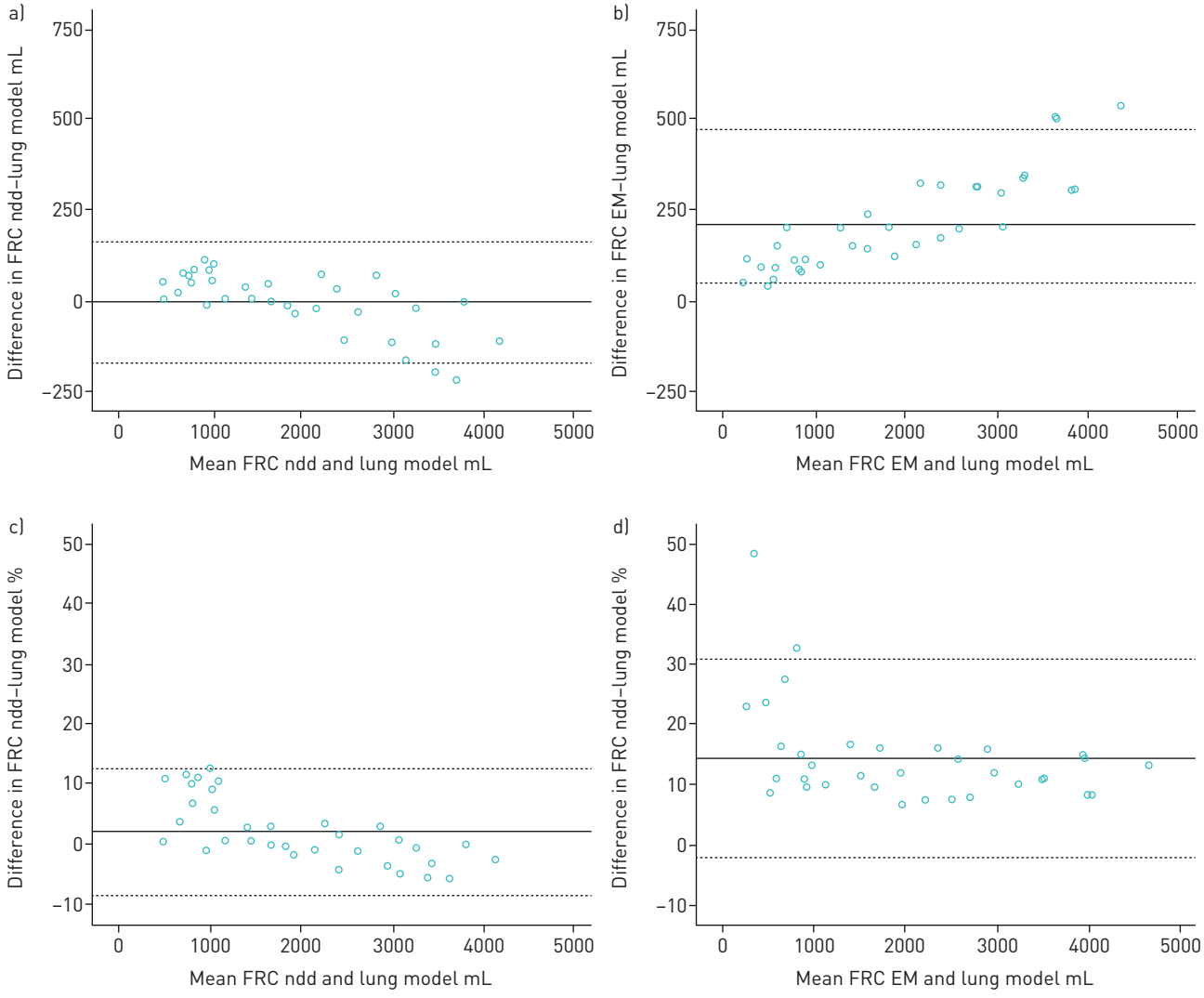

FIGURE 2 Bland-Altman plots of *in vitro* MBNW FRC measurements of a) ndd (EasyOne Pro Lab) and lung model and b) EM (Exhalyzer D) and lung model in absolute values and c) ndd and lung model and d) EM and lung model in relative values. Data are plotted as measured FRC minus lung model FRC, expressed as absolute and relative difference versus mean of measured and lung model FRC values. The middle line represents the mean difference and the upper and lower (dashed) lines the upper and lower limits of agreement (mean difference ± 1.96 SD).

TABLE 1 Mean difference between measured FRC and FRC lung model

FRC lung model		ndd	EM
All	Absolute	−4 (−34–26)	212 (166–259)
	Relative	2.1 (0.2–3.9)	14.3 (11.4–17.2)
<1000 mL	Absolute	59 (35–84)	100 (71–128)
	Relative	7.6 (4.7–10.5)	20.0 (12.5–27.5)
>1000 mL	Absolute	−39 (−76–−2)	274 (219–329)
	Relative	−1.0 (−2.2–0.3)	11.2 (9.8–12.6)

Data are presented as mean (95% CI). Absolute measurements are in millilitres, relative in percentage. FRC: functional residual capacity; ndd: EasyOne Pro LAB; EM: Exhalyzer D.

fixed tidal breathing protocol and was excluded. After exclusion, a total of 38 participants were eligible for analysis, their characteristics are outlined in table 2.

Comparison of ndd and EM: FRC, LCI and secondary outcome parameters

FRC_{ndd} was significantly lower than FRC_{EM}, mean difference −11.8% ($p \leqslant 0.01$, limits of agreement −25.6 to 2%) (figure 3, table 3).

The LCI was comparable between the two devices, mean difference −1.3% ($p=0.31$, limits of agreement −21.9 to 19.3%) (figure 3). LCI differed >10% in seven (39%) children and five (25%) adults between the two devices. Within-test repeatability of LCI was good in both devices; mean coefficient of variation (SD) 4.2 (3.1) using ndd and 3.8 (2.7) using EM.

Cumulative expired volume (CEV) was lower for both, children and adults when measured by ndd in comparison to EM. Respiratory rate was significantly higher using ndd in both groups, V_T was significantly lower in children on ndd compared to EM and a trend was seen towards a lower $V_T \cdot kg^{-1}$ in children on ndd (table 3).

Comparison of FRC measurements between MBNW, body plethysmography and helium dilution technique

FRC_{ndd} was significantly lower than FRC_{pleth} with a mean difference of −7.0% ($p \leqslant 0.01$ limits of agreement −23.2 to 9.2%) (table 4, figure S1).

Mean difference between FRC_{EM} and FRC_{pleth} was 5.7% ($p \leqslant 0.01$, limits of agreement −11.2 to 22.6%) (figure S1). FRC_{EM} did not significantly differ from FRC_{pleth} in children in contrast to adults (table 3).

Mean FRC_{ndd} and FRC_{He} did not significantly differ, mean difference 2.1% ($p=0.46$, limits of agreement −17.9 to 22.1%) (table 3. figure S1). FRC_{EM} was significantly higher compared to $FRC_{He,}$ mean difference 16.3% ($p \leqslant 0.01$, limits of agreement −10.6 to 43.2) (figure 3, table 3).

TABLE 2 Patient characteristics

	Child (n=18)	Adult (n=20)
Age years	12.5 (6.1–17.1)	37.5 (18.8–64.9)
Male	7 (38.9)	12 (60)
Height cm	156.8 (124–184)	179.9 (167–192)
Weight kg	47.6 (25–78)	74.1 (53–103)
FEV_1	0.08 (−2.0–1.5)	0.17 (−1.6–1.3)
FVC	0.13 (−2.2–1.4)	0.17 (−1.4–1.7)
FEV_1/FVC	−0.16 (−2.0–1.2)	−0.07 (−1.6–1.3)
MEF_{25}	−0.07 (−1.2–1.2)	0.24 (−1.3–1.4)
FRC_{pleth} L	2.1 (1.2–4.2)	3.8 (2.8–5.6)
FRC_{He} L[#]	1.9 (1.0–3.8)	3.6 (2.7–4.7)

Data are presented as mean (range) or n (%). FEV_1, FVC, FEV_1/FVC and MEF_{25} are expressed as z scores. [#]: n=17; 1 child was not able to perform He test due to technical problems. FEV_1: forced expiratory volume in 1 s; FVC: forced vital capacity; MEF_{25}: maximal expiratory flow at 25% of vital capacity; FRC: functional residual capacity; pleth: body plethysmography; He: helium dilution technique.

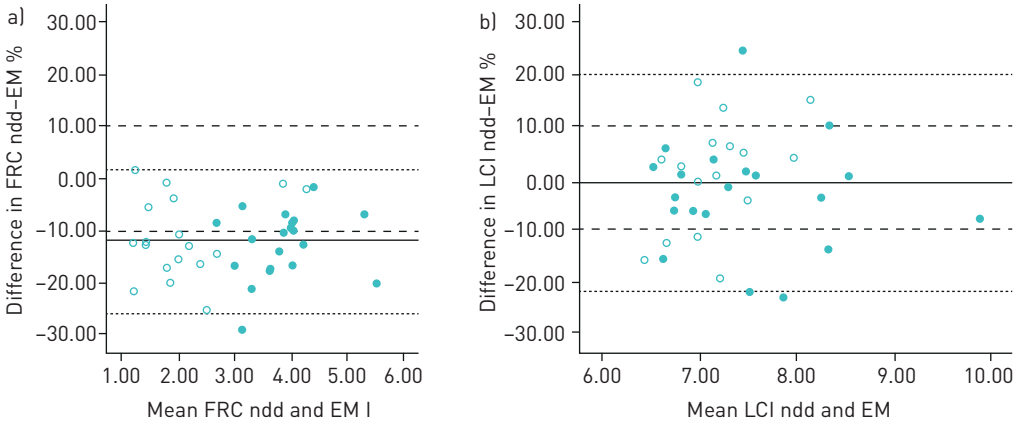

FIGURE 3 Bland-Altman plots of a) FRC (functional residual capacity) and b) LCI (lung clearance index) agreement between ndd (EasyOne Pro Lab) and EM (Exhalyzer D). Data are plotted as relative differences (in %).The continuous line represents the mean difference, the fine dashed lines the upper and lower limits of agreement (mean difference ± 1.96 SD) and the rough dashed lines the defined limits of clinical acceptability (10 to −10%). Open circles represent children; closed circles adults.

Relaxed versus *fixed tidal breathing*

Mean FRC and LCI (p-value, limits of agreement) were −1.8% (p=0.45, −32.1 to 28.6%) and −1% (p=0.37, −22.5 to 20.5%) using ndd and −1.4%(p=0.33, 27.8 to 25.0%) and −1.5 (p=0,17, −18.7 to 15.7%) using EM, respectively (figure S2). Overall, 11 (61%, ndd) and 10 (56%, EM) of the children and 6 (30%, ndd) and 6 (30%, EM) of the adults had a relaxed tidal breathing pattern that fulfilled the ATS/ERS criteria for fixed tidal breathing. In children, only a significant difference was seen in a lower respiratory rate on EM during fixed tidal breathing (see Table S1 online supplementary materials). In adults, respiratory rate was significantly lower on ndd during fixed tidal breathing. V_T and $V_T \cdot kg^{-1}$ were significantly higher on both devices during fixed tidal breathing (see Table S1 online supplementary materials).

Discussion

Our study demonstrates that for MBNW the ndd and EM cannot be used interchangeably in healthy children and adults. FRC_{EM} yielded higher values compared to FRC_{ndd}, both, *in vitro* and *in vivo*. The LCI was comparable between the two devices; however, there were differences within subjects in almost one-third of measurements. Moreover, both devices showed deviations in FRC measurement *in vitro* and *in vivo* compared to pleth as well as the helium dilution technique. FRC_{ndd} was lower compared to FRC_{pleth} in 38% of subjects; FRC_{EM} was higher than FRC_{pleth} and FRC_{He} in 50 and 66% of subjects

TABLE 3 Comparison of FRC between the different test methods

	Absolute (l)		Relative (%)		p-value
FRC					
ndd *versus* pleth					
Child	−0.16	(−0.24; −0.07)	−7.7	(−12.1; −3.4)	**<0.005**
Adult	−0.24	(−0.39; −0.09)	−6.3	(−10.2; −2.4)	**0.04**
ndd *versus* He					
Child	0.07	(−0.03; 0.17)	3.9	(−1.3; 9.1)	0.15
Adult	0.01	(−0.17; 0.19)	0.5	(−4.3; 5.3)	0.88
EM *versus* pleth					
Child	0.08	(−0.03; 0.18)	4.0	(−0.6; 8.5)	0.15
Adult	0.27	(0.13; 0.42)	7.1	(3.3; 11.0)	**<0.005**
EM *versus* He					
Child	0.31	(0.19; 0.43)	17.4	(10.3; 24.5)	**0.04**
Adult	0.52	(0.29; 0.75)	15.3	(8.8; 21.7)	**<0.005**

Data are presented as mean (95% CI). The p-value was calculated based on the absolute values of FRC_{MBNW} *versus* FRC_{pleth} or FRC_{He}. FRC: functional residual capacity; ndd: EasyOne Pro LAB; pleth: body plethysmography; He: helium dilution technique; EM: Exhalyzer D.

TABLE 4 Comparison of outcome parameters between EasyOne Pro LAB (ndd) and Exhalyzer D (EM) during relaxed tidal breathing

	ndd		EM		p-value
FRC L					
Child	1.9	(1.5; 2.4)	2.2	(1.7; 2.6)	<0.005
Adult	3.6	(3.3; 3.9)	4.1	(3.8; 4.4)	<0.005
LCI					
Child	7.2	(6.8; 7.5)	7.1	(6.9; 7.3)	0.83
Adult	7.3	(6.9; 7.8)	7.6	(7.1; 8.1)	0.14
CEV L					
Child	13.9	(10.7; 17.0)	16.6	(13.6; 19.7)	<0.005
Adult	26.3	(23.6; 29.0)	33.5	(29.5; 37.4)	<0.005
V_T mL					
Child	543	(452; 633)	653	(513; 792)	0.04
Adult	862	(717; 1006)	840	(756; 926)	0.63
$V_T \cdot kg^{-1}$ mL·kg^{-1}					
Child	11.6	(10.3; 13.0)	15.4	(11.5; 16.0)	0.06
Adult	11.6	(9.9; 13.4)	11.4	(10.4; 12.4)	0.72
RR min^{-1}					
Child	22.5	(19.8; 25.2)	15.4	(13.0; 17.8)	<0.005
Adult	15.3	(13.1; 17.5)	11.9	(10.4; 13.3)	<0.005

Data are presented as mean (95% CI), unless otherwise stated. FRC: functional residual capacity; LCI: lung clearance index; CEV: cumulative expired volume; V_T: tidal volume; RR: respiratory rate.

respectively. In addition, this study shows that predefined breathing patterns have no significant effect on FRC_{MBNW} and LCI.

In vitro

Neither of the two devices fulfilled the ATS/ERS criteria of >95% of measurements within 5% of $FRC_{lungmodel}$. In contrast to published data, FRC_{EM} was higher than the defined accuracy rate of 5% in all tests [13, 17, 20].

In theory, several factors could explain the differences between FRC_{MBNW} and $FRC_{lungmodel}$. First, the volume of dead space and correction for dead space are different between the two devices. The ndd device has a larger dead space area than EM. FRC is automatically corrected for dead space in both devices. A study investigating the effects of dead space showed that increasing dead space did not have a significant effect on FRC in healthy adults [31]. This suggests that it is unlikely that dead space differences influenced our results. Second, CO_2 in exhaled breath might be crucial for accurate calculation of molar mass with the ultrasonic sensor utilised for MBNW in these devices. Unfortunately, simulation of physiological CO_2 concentrations during expiration of human breathing patterns is not yet possible in the lung model. One can either choose between no CO_2 administration or continuous CO_2. It is unlikely that continuous CO_2 would affect the comparison of both devices using the lung model.

Third, $FRC_{lungmodel}$ was based on the volume of the lung model when the MBNW device was connected. Connection of the devices led to a reduction in $FRC_{lungmodel}$, which can be explained by the internal resistance of the devices. This reduction was more pronounced in the EM device compared to ndd. Previous studies did not describe whether the FRC was measured with or without connection of the MBNW device [12, 13, 17, 20]. Interestingly, a relative volume-dependent difference was seen in $FRC_{lungmodel}$ <1000 mL in both devices, suggesting that the devices are less suitable for lower lung volumes. This needs to be explored in further research.

In vivo

FRC_{EM} was significantly higher than FRC_{ndd} and not within the limits of clinical acceptability, indicating that the two devices are not interchangeable. The underestimation of FRC_{ndd}, and the overestimation of FRC_{EM} was consistent *in vitro* and *in vivo* and in other studies that investigated the two devices [14, 16, 18, 20].

In agreement with other studies, the EM yielded a slightly overestimated FRC for the group compared to FRC_{pleth}; however, in 50% of subjects measured FRC was overestimated more than can be considered clinically acceptable [14, 15, 17, 18, 20]. In contrast the ndd does underestimate the FRC compared to FRC_{pleth}. These data suggest a systematic measurement error within both devices.

In the ndd, the underestimation of FRC was clearly reduced using the new software applying the N_2 algorithm, compared to studies that used older software versions [16–18, 20].

Our data are in agreement with the retrospective data of Tonga *et al.* [17], who also found an underestimation of FRC_{ndd} *versus* FRC_{pleth} of around −10% using the new software. Software changes led to the reduction of underestimation of FRC, by improving delay time correction between flow and gas measurement points and the switch of use of a prototype expirogram derived from the early breaths from the washout to compute N_2 to a point-by-point measurement during the entire expirogram. The change in delay time potentially has a large effect on FRC calculation [32]. However, in almost 40% of the subjects the underestimation of FRC was still not clinically acceptable, which implies the measurement can be further improved.

The mean LCI for the group was comparable, in contrast to previous studies that compared the two devices and showed a significantly higher LCI measured by EM compared to ndd [14, 20]. Nonetheless, in the present dataset there was a wide range and the limits of a clinically acceptable difference of <10% were not met in almost one-third of subjects, as was previously shown by Poncin *et al.* [14] using an old software of ndd (version 3.37) [18]. Since the LCI is calculated as CEV/FRC, it can be expected that differences in CEV and/or FRC will explain the variability in LCI. Although CEV and FRC measured by EM were consistently higher than ndd, these parameters did not explain the variation in LCI, in contrast to the findings of Poncin *et al.* [14].

In contrast to the absence of effect of dead space on the FRC, the difference in dead space correction could have had an effect on LCI in our *in vivo* measurements [31]. For EM LCI is automatically corrected for the dead space component. This is not the case for ndd, although this option became available in the latest software version as LCI at airway opening (LCI_{ao}) (version V3.05.01.07, 03-2019). A *post hoc* recalculation in $LCI_{ao\text{-}ndd}$ led to a significantly lower LCI compared to LCI_{EM}, mean difference −11.8% ($p<0.005$, limits of agreement −24.32 to 1.6) and thus led to a decrease in comparability between the two devices.

Breathing pattern

The use of a fixed and relaxed tidal breathing protocol did not result in any systematic difference in FRC or LCI, in contrast to other studies investigating the influence of breathing pattern that showed inconsistent effects on FRC and LCI [24, 25]. The explanation might be that many of our subjects fulfilled the ATS/ERS criteria of fixed tidal breathing during relaxed tidal breathing. Many of our smaller adult subjects experienced difficulties to fulfil the ATS/ERS fixed tidal breathing criteria [11]. Therefore, we advise a standard breathing protocol with relaxed tidal breathing pattern with a minimal breath size of around 8 $mL \cdot kg^{-1}$ to overcome dead space ventilation.

Changes in breathing pattern, different from the change between a fixed and relaxed tidal breathing protocol, can influence MBNW outcome parameters [11, 24, 25]. Although we did not perform a systematic evaluation of the experience of subjects on both devices, issues that might have influenced relaxed tidal breathing pattern were raised spontaneously by the subjects. For the EM, a feeling of resistance during breathing was reported. This is supported by the previously mentioned reduction of $FRC_{lungmodel}$ during connection of the EM *in vitro*. For ndd, some subjects felt "rushed" by the on-demand oxygen supply, which explains the higher respiratory rate for ndd compared to EM. Improvement could be reached by changing from on demand to a continuous oxygen supply. The effect of these issues would be interesting to explore in future research. Still, if a change of breathing pattern were the explanation for the differences between the devices, we would have expected the *in vitro* measurements to be more consistent, as the effect of a different breathing pattern was excluded using a standardised ventilator.

Strengths and limitations

This is the first prospective comparative study assessing two MBNW devices both *in vitro* and *in vivo* over a broad FRC range (FRC_{pleth} varying between 1.2 and 5.6 L) including comparison with plethysmography and helium dilution for both devices.

A limitation of our study is that the power calculation was based on the study of Raaijmakers *et al.* [20]; at that time the only available study on this topic. According to the data published by Tonga *et al.* [17], 34 volunteers would have been needed for each group to achieve a power of 80% with a two-sided significance level (α) of 0.05. Therefore, it may be argued that the present study is under powered [14, 17]. Nevertheless, we were able to detect clinically relevant and statistically significant differences.

Conclusion

Numerous studies demonstrated that MBNW is an important tool for assessing early changes in peripheral airways in CF [4, 7, 10, 33, 34]. There is growing interest to support the use of LCI as primary outcome

parameter in interventional and clinical studies [35–37]. Our study shows that there is room for further improvement of available equipment, given the scientific importance of addressing small airways disease. Based on the present study there is no clear preference for one of the two devices; however, it is essential to choose one device for clinical follow-up of patients as well as in longitudinal research protocols.

Acknowledgements: The authors would like to thank the healthy control subjects for their participation in the study, as well as statistician Judith Vonk (University of Groningen, University Medical Center Groningen, Groningen Research Institute for Asthma and COPD, and Dept of Epidemiology, Groningen, the Netherlands) for her help.

Conflict of interest: A.M. Zwitserloot reports a grant from Stichting Astma Bestrijding and that ndd Medical Technologies provided the EasyOne Pro LAB MBNW device and lung model setup for this study; and money to her institution for an advisory board for GSK outside the submitted work. E.J. van den Berg has nothing to disclose. L.H.A. Raaijmakers has nothing to disclose. W.E. Stalman has nothing to disclose. M. van Smaalen has nothing to disclose. M. van den Berge reports research grants to his university from GlaxoSmithKline, TEVA, Chiesi and Novartis, outside the submitted work. M. Gappa has nothing to disclose. G.H. Koppelman reports grants from the Lung Foundation of the Netherlands during the conduct of the study, TETRI Foundation, UBBO EMMIUS Foundation, TEVA the Netherlands, VERTEX and GSK, and that he has participated in an advisory board for GSK, outside the submitted work. B.W.M. Willemse has nothing to disclose.

Support statement: This research project was funded by from the Stichting Astma Bestrijding, the Netherlands, under grant agreement 2014/062. The funders had no role in the study design, results or interpretation. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Saglani S, Payne DN, Zhu J, *et al.* Early detection of airway wall remodeling and eosinophilic inflammation in preschool wheezers. *Am J Respir Crit Care Med* 2007; 176: 858–864.
2 Johnson JR, Hamid Q. Appraising the small airways in asthma. *Curr Opin Pulm Med* 2012; 18: 23–28.
3 Tiddens HA, Donaldson SH, Rosenfeld M, *et al.* Cystic fibrosis lung disease starts in the small airways: can we treat it more effectively? *Pediatr Pulmonol* 2010; 45: 107–117.
4 Aurora P, Gustafsson P, Bush A, *et al.* Multiple breath inert gas washout as a measure of ventilation distribution in children with cystic fibrosis. *Thorax* 2004; 59: 1068–1073.
5 Fuchs SI, Gappa M, Eder J, *et al.* Tracking lung clearance index and chest CT in mild cystic fibrosis lung disease over a period of three years. *Respir Med* 2014; 108: 865–874.
6 Hjalmarson O, Sandberg K. Abnormal lung function in healthy preterm infants. *Am J Respir Crit Care Med* 2002; 165: 83–87.
7 Kozlowska WJ, Bush A, Wade A, *et al.* Lung function from infancy to the preschool years after clinical diagnosis of cystic fibrosis. *Am J Respir Crit Care Med* 2008; 178: 42–49.
8 Long FR, Williams RS, Castile RG. Structural airway abnormalities in infants and young children with cystic fibrosis. *J Pediatr* 2004; 144: 154–161.
9 Fuchs SI, Ellemunter H, Eder J, *et al.* Feasibility and variability of measuring the lung clearance index in a multi-center setting. *Pediatr Pulmonol* 2012; 47: 649–657.
10 Gustafsson PM, Aurora P, Lindblad A. Evaluation of ventilation maldistribution as an early indicator of lung disease in children with cystic fibrosis. *Eur Respir J* 2003; 22: 972–979.
11 Robinson PD, Latzin P, Verbanck S, *et al.* Consensus statement for inert gas washout measurement using multiple- and single- breath tests. *Eur Respir J* 2013; 41: 507–522.
12 Fuchs SI, Buess C, Gappa M. In vitro validation of multiple breath washout using ultrasonic equipment. *Eur Respir J* 2011; Suppl. 55, 38: 355s.
13 Singer F, Houltz B, Latzin P, *et al.* A realistic validation study of a new nitrogen multiple-breath washout system. *PLoS ONE* 2012; 7: e36083.
14 Poncin W, Singer F, Aubriot AS, *et al.* Agreement between multiple-breath nitrogen washout systems in children and adults. *J Cyst Fibros* 2017; 16: 258–266.
15 Jensen R, Stanojevic S, Gibney K, *et al.* Multiple breath nitrogen washout: a feasible alternative to mass spectrometry. *PLoS ONE* 2013; 8: e56868.
16 Zwitserloot AM, Raaijmakers LHA, Fuchs SI, *et al.* Comparison of FRC and TLC measured in whole body plethysmography and multiple breath washout in healthy adult volunteers. *Am J Respir Crit Care Med* 2014; 189: A2988.
17 Tonga KO, Robinson PD, Farah CS, *et al.* In vitro and in vivo functional residual capacity comparisons between multiple-breath nitrogen washout devices. *ERJ Open Res* 2017; 3: 00011-2017.
18 Clermons K, Cornelussen S, Martin L, *et al.* Selecting a multiple-breath N_2 washout device: the search continues. *Eur Respir J* 2015; 46: Suppl. 59, PA1581.
19 Clermons K, Cornelussen S, Martin L, *et al.* Helium or nitrogen: that's the question? *Eur Respir J* 2014; 44: Suppl. 58, P1993.
20 Raaijmakers L, Jensen R, Stanojevic S, *et al.* Validation of multiple breath washout devices. *J Cyst Fibros* 2017; 16: e22–e23.
21 Verbanck S, Thompson BR, Schuermans D, *et al.* Ventilation heterogeneity in the acinar and conductive zones of the normal ageing lung. *Thorax* 2012; 67: 789–795.
22 Aurora P, Kozlowska W, Stocks J. Gas mixing efficiency from birth to adulthood measured by multiple-breath washout. *Respir Physiol Neurobiol* 2005; 148: 125–139.
23 Crawford AB, Makowska M, Paiva M, *et al.* Convection- and diffusion-dependent ventilation maldistribution in normal subjects. *J Appl Physiol* 1985; 59: 838–846.
24 Gronkvist M, Bergsten E, Gustafsson PM. Effects of body posture and tidal volume on inter- and intraregional ventilation distribution in healthy men. *J Appl Physiol* 2002; 92: 634–642.

25 Yammine S, Singer F, Gustafsson P, *et al.* Impact of different breathing protocols on multiple-breath washout outcomes in children. *J Cyst Fibros* 2014; 13: 190–197.
26 Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.
27 Wanger J, Clausen JL, Coates A, *et al.* Standardisation of the measurement of lung volumes. *Eur Respir J* 2005; 26: 511–522.
28 Cliff IJ, Evans AH, Pantin CF, *et al.* Comparison of two new methods for the measurement of lung volumes with two standard methods. *Thorax* 1999; 54: 329–333.
29 Brown R, Leith DE, Enright PL. Multiple breath helium dilution measurement of lung volumes in adults. *Eur Respir J* 1998; 11: 246–255.
30 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307–310.
31 Benseler A, Stanojevic S, Jensen R, *et al.* Effect of equipment dead space on multiple breath washout measures. *Respirology* 2015; 20: 459–466.
32 Horsley A, Macleod K, Gupta R, *et al.* Enhanced photoacoustic gas analyser response time and impact on accuracy at fast ventilation rates during multiple breath washout. *PLoS One* 2014; 9: e98487.
33 Kraemer R, Blum A, Schibler A, *et al.* Ventilation inhomogeneities in relation to standard lung function in patients with cystic fibrosis. *Am J Respir Crit Care Med* 2005; 171: 371–378.
34 Hoo AF, Thia LP, Nguyen TT, *et al.* Lung function is abnormal in 3-month-old infants with cystic fibrosis diagnosed by newborn screening. *Thorax* 2012; 67: 874–881.
35 Subbarao P, Milla C, Aurora P, *et al.* Multiple-breath washout as a lung function test in cystic fibrosis. A cystic fibrosis foundation workshop report. *Ann Am Thorac Soc* 2015; 12: 932–939.
36 Amin R, Stanojevic S, Kane M, *et al.* A randomized controlled trial to evaluate the lung clearance index as an outcome measure for early phase studies in patients with cystic fibrosis. *Respir Med* 2016; 112: 59–64.
37 Kent L, Reix P, Innes JA, *et al.* Lung clearance index: Evidence for use in clinical trials in cystic fibrosis. *J Cyst Fibros* 2014; 13: 123–138.