




In vitro and *in vivo* functional residual capacity comparisons between multiple-breath nitrogen washout devices

Katrina O. Tonga^{1,2,3,4}, Paul D. Robinson ^{1,2,5}, Claude S. Farah^{1,2,4},
Greg G. King^{1,2,3} and Cindy Thamrin^{1,2}

Affiliations: ¹Woolcock Institute of Medical Research, University of Sydney, Sydney, Australia. ²Sydney Medical School, University of Sydney, Sydney, Australia. ³Dept of Respiratory Medicine, Royal North Shore Hospital, Sydney, Australia. ⁴Dept of Respiratory Medicine, Concord Hospital, Sydney, Australia. ⁵Dept of Respiratory Medicine, The Children's Hospital at Westmead, Sydney, Australia.

Correspondence: Katrina O. Tonga, Woolcock Institute of Medical Research, University of Sydney, Sydney, NSW 2037, Australia. E-mail: katrina.tonga@sydney.edu.au

ABSTRACT Functional residual capacity (FRC) accuracy is essential for deriving multiple-breath nitrogen washout (MBNW) indices, and is the basis for device validation. Few studies have compared existing MBNW devices. We evaluated *in vitro* and *in vivo* FRC using two commercial MBNW devices, the Exhalizer D (EM) and the EasyOne Pro LAB (ndd), and an in-house device (Woolcock in-house device, WIMR).

FRC measurements were performed using a novel syringe-based lung model and in adults (20 healthy and nine with asthma), followed by plethysmography (FRC_{pleth}). The data were analysed using device-specific software. Following the results seen with ndd, we also compared its standard clinical software (ndd v.2.00) with a recent upgrade (ndd v.2.01).

WIMR and EM fulfilled formal *in vitro* FRC validation recommendations (>95% of FRC within 5% of known volume). Ndd v.2.00 underestimated *in vitro* FRC by >20%. Reanalysis using ndd v.2.01 reduced this to 11%, with 36% of measurements ≤5%. *In vivo* differences from FRC_{pleth} (mean±SD) were 4.4±13.1%, 3.3±11.8%, -20.6±11% (p<0.0001) and -10.5±10.9% (p=0.005) using WIMR, EM, ndd v.2.00 and ndd v.2.01, respectively.

Direct device comparison highlighted important differences in measurement accuracy. FRC discrepancies between devices were larger *in vivo*, compared to *in vitro* results; however, the pattern of difference was similar. These results represent progress in ongoing standardisation efforts.



@ERSpublications

Multiple-breath washout devices are not yet comparable <http://ow.ly/bB7b30eAs0c>

Cite this article as: 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 [<https://doi.org/10.1183/23120541.00011-2017>].



Introduction

Abnormalities in the small airways can be detected using the multiple-breath nitrogen washout (MBNW) test, which provides a more sensitive measure of small airway function than spirometry [1–3]. MBNW provides insightful information about the small airways, using indices that assess gas-mixing efficiency in the lung (Lung Clearance Index, LCI) and mechanistic information about ventilation heterogeneity in the conduction and diffusion-dependent airways (Scond and Sacin). These indices rely on accurate functional residual capacity (FRC) measurements [4].

Advances in technology have allowed the development of new commercial MBNW devices; however, there are still limited published data comparing measurements between devices. Previous studies have compared in-house and/or commercial devices against standard body plethysmography as well as mass spectrometry, often considered the gold standard [5–8]. Many of these studies have used lung models, assessed measurements in the paediatric population and used inert gases rather than N₂ [9–11]. Limited published data have also shown inconsistencies between lung model and adult measurements [6].

Furthermore, the effect of software upgrades in a rapidly evolving field has not been studied extensively. Previous studies in one specific device have shown that software changes can have a significant impact on results [12, 13]. Recommendations for MBNW techniques have been published in the European Respiratory Society/American Thoracic Society (ERS/ATS) Consensus statement [4]; however, ongoing work is still needed to improve the standardisation of equipment, technical specifications and algorithms for the calculation of indices.

The aim of this study was to evaluate FRC measured using two commercial MBNW devices, the Exhalyzer D (EM) from Eco Medics AG (Duernten, Switzerland) and the EasyOne Pro LAB (nnd) from nnd Medical Technologies (Zurich, Switzerland). We also examined a third, independent, previously published [2] in-house device at the Woolcock Institute of Medical Research (WIMR), which measures N₂ directly, in contrast to the commercial devices. FRC measurements from the devices were compared against a known volume using a syringe lung model (*in vitro*) and against plethysmographic lung volume in healthy and asthmatic adult subjects (*in vivo*). In addition, we compared analyses using two different software versions of the nnd device, because a software update involving major changes to N₂ calculation has recently become available. Previous data [14, 15] have shown significant FRC underestimation using the older, widely available software.

Methods

Study design

In vitro and *in vivo* FRC measurements were performed using the three different MBNW devices, in random order as determined by a computer-generated randomisation sequence. The study was approved by the Northern Sydney Local Health District Human Research Ethics Committee (protocol no. LNR/16/HAWKE/11). Written informed consent was obtained from all recruited participants.

The lung model

In vitro measurements were performed using an optimised syringe lung model (figure 1). A 3 L volume calibration Hans Rudolph syringe (5530 series) was modified to produce the physiological expirogram



FIGURE 1 Physical lung model composed of a) a 3 L Hans Rudolph syringe, b) an attachment made from 18 flexible tubes, c) a helical mixer device inserted at the syringe entrance, and d) an adjustable syringe stopper.

encountered during *in vivo* testing. This was accomplished by incorporating an attachment on the front of the syringe consisting of 18 flexible Tygon (S3 E-3603) tubes of varying lengths (internal diameter of 0.048 cm and lengths ranging from 10 to 49.5 cm) and permeability coefficients for CO₂ of $360 \times 10^{-11} \text{ cm}^2 \cdot \text{s}^{-1} \cdot \text{cmHg}^{-1}$, N₂ of $40 \times 10^{-11} \text{ cm}^2 \cdot \text{s}^{-1} \cdot \text{cmHg}^{-1}$ and O₂ of $80 \times 10^{-11} \text{ cm}^2 \cdot \text{s}^{-1} \cdot \text{cmHg}^{-1}$, to produce the phase I and II portions of the expirogram, and a 3D-printed helical mixer device inserted at the syringe entrance to optimise gas mixing and produce a smooth phase III.

The target FRC was calculated by adding the known syringe volume to the dead space of the lung model attachments. The syringe volume was adjusted via a stopper on the syringe plunger to predetermined positions. The dead space of the attachments was 0.310 L, determined from the computer-aided design specifications and confirmed with water displacement.

In vitro study

In vitro measurements were performed in triplicate on each device using four different FRC volumes (1.51 L, 1.81 L, 2.11 L and 2.31 L). A standardised adult protocol (tidal volume of 1–1.3 L) [2, 4, 16–18] was used. After at least 5 syringe strokes with a stable end expiratory volume, the washout phase was commenced and 100% oxygen was switched on. Syringe strokes were continued until end-tidal N₂ concentration decreased to 1/40th of the starting end-tidal N₂ concentration [4]. Between measurements, at least 10 strokes were first performed to expel any residual oxygen within the syringe lung model and to ensure the N₂ had returned to baseline. A single operator (K.O. Tonga) performed all measurements under ambient temperature and pressure dry (ATPD) conditions.

In vivo study

Healthy volunteers were recruited from the WIMR and defined as current non-smokers with <10 pack-years smoking history and with no history of acute respiratory illness within the preceding month. Subjects with asthma were recruited if they had a physician diagnosis of asthma: history of asthma symptoms, previously documented significant bronchodilator reversibility on spirometry and/or a positive bronchoprovocation challenge test and on inhaled bronchodilator and/or inhaled corticosteroid medication. Short acting β -agonists were withheld for 6 h and long-acting β -agonists for 24 h before testing.

All participants were over the age of 18 years and completed a standardised interview on respiratory and general health before performing pulmonary function tests (PFTs) in the following order: MBNW, spirometry and body plethysmography. All tests were performed in a single session at the Woolcock Institute of Medical Research.

Spirometry and body plethysmography were performed according to ATS/ERS Guidelines, using a BodyBox 5500 (Medisoft Corporation, Sorlines, Belgium). MBNW was performed in triplicate in the seated position, using a noseclip and device-specific bacterial filter and mouthpiece attachments. Tests were conducted according to ERS/ATS consensus statement [4], using a standardised adult protocol (tidal volume of 1–1.3 L) [2, 16–18]. After at least 5 breaths with a stable end expiratory lung volume, the washout phase was commenced and 100% oxygen was switched on. Breaths were continued until end-tidal N₂ concentration decreased to 1/40th of the starting end-tidal N₂ concentration [4]. The time interval between measurements was standardised to twice the previous washout time [4].

MBNW hardware and software

Details of the WIMR, EM and ndd devices are provided in the supplementary material. Briefly, the WIMR device measures N₂ directly via a side-stream N₂ analyser [2]. The commercial devices both measure N₂ indirectly, based on side-stream CO₂ and O₂ in the EM device and molar mass and CO₂ in the ndd device [6]. Daily calibration and/or verification of each device was performed as described in the supplementary material.

FRC was calculated as the net volume of N₂ expired divided by the difference between end-tidal N₂ concentration at the start and end of the washout portion of the test [4]. FRC values were corrected for the pre-capillary dead space volume for each device. MBNW data were analysed using software specific to each device. For the WIMR device, custom-written software (Solver Version 1.3.2.18) was used. For the EM device, Spiroware Version 3.1.6 was used. For the ndd device, the same measurements were analysed using two different software versions (described in the supplementary material): clinical software Version 2.00.01.05 (termed “ndd v.2.00”) and the recent upgrade released in March 2016, Version 2.01.00.09 (termed “ndd v.2.01”). The ndd v.2.01 software contains updates to the N₂ calculation method and improved flow-gas delay synchronisation.

Measurements were included for analyses if acceptability criteria were met. For the *in vitro* study, measurements were deemed acceptable if three tests were within 10% of the mean FRC [19] value across the triplicate measures. For the *in vivo* study, measurements were deemed acceptable if two or more tests were performed adequately according to the ERS/ATS Consensus statement [4]. Tests were excluded if there was evidence of leak during testing, the tidal volume of the first breath or more than a third of breaths during washout was outside 1–1.3 L, and if the end-tidal N₂ concentration did not reach the recommended 1/40th of the initial concentration during data acquisition.

Statistical analysis

FRC data were analysed using IBM SPSS Version 22, and graphs were generated using GraphPad Prism Version 6.0. Summary data are presented as mean±SD. The accuracy of *in vitro* FRC was assessed according to the consensus statement [4] (FRC values within 5% of the known volume for at least 95% of values) and expressed as absolute (L) difference (measured FRC – lung model FRC) and relative (%) difference (absolute difference×100/lung model FRC). *In vivo* FRC was compared with body plethysmography FRC (FRC_{pleth}). FRC differences between devices were assessed using Bland and Altman plots with 95% limits of agreement and by non-parametric one-way repeated measures ANOVA (Friedman's test) and *post hoc tests*. p-values <0.05 were considered statistically significant.

Results

In vitro comparison

A total of 108 measurements were performed across the three MBNW devices (table 1). Differences between measured and lung model FRC values are shown in figure 2 and table 2. FRC accuracy was within the specified 5% accuracy range of the lung model FRC for 100% of WIMR measurements and 97% of EM measurements. All FRC measurements using ndd v.2.00 were underestimated and not within the specified 5% accuracy range. Using ndd v.2.01, accuracy improved, although only 36% of measurements were within the 5% accuracy range.

In vivo comparison

A total of 29 subjects (20 healthy controls and 9 asthmatics) were included in the analyses, and their characteristics are outlined in table 3. The mean (±SD) FRC measured by the WIMR (3.27±0.82 L) and EM (3.56±0.92 L) devices did not differ significantly from FRC_{pleth} (3.44±0.77 L) or from each other, and FRC measured by ndd v.2.00 (2.71±0.64 L) was significantly lower than FRC_{pleth}, WIMR and EM (p<0.0001) (figure 3 and table 4). The same pattern of FRC differences was observed in healthy control and asthmatic subjects.

When the ndd data were reanalysed using ndd v.2.01 software, end-tidal N₂ concentrations became systematically higher, such that only 9 of the 29 subjects had measurements that reached an acceptable end-tidal N₂ concentration. However, a majority of the measurements exhibited a stable plateau at the end of the washout. For the purposes of this study, all 29 subjects were included for analysis to allow comparison of the effect on FRC. FRC increased when reanalysed using ndd v.2.01 (3.06±0.71 L) and was also significantly lower than FRC_{pleth} (p=0.011) and the EM device (p<0.0001); however, it did not differ from the WIMR device.

Discussion

Summary of results

This is the first study to compare *in vitro* and *in vivo* FRC measurements in healthy and asthmatic adults, using two currently available commercial devices and one in-house MBNW device. The WIMR device

TABLE 1 *In vitro* functional residual capacity (FRC) measurements on each device

Lung model FRC volume	Trials	WIMR	EM	ndd v.2.00	ndd v.2.01
2.31	9	2.30±0.02	2.39±0.02	1.78±0.14	2.16±0.04
2.11	9	2.10±0.02	2.17±0.02	1.72±0.12	2.00±0.03
1.81	9	1.84±0.02	1.88±0.02	1.37±0.11	1.72±0.03
1.51	9	1.53±0.03	1.53±0.02	1.25±0.05	1.43±0.02

Data presented as mean±SD, unless otherwise stated. Measurements are in litres. WIMR: Woolcock in-house device; EM: Exhalyzer D device; ndd v.2.00: EasyOne Pro LAB device Version 2.00.01.05; ndd v.2.01: EasyOne Pro LAB device analysed in Version 2.01.00.09.

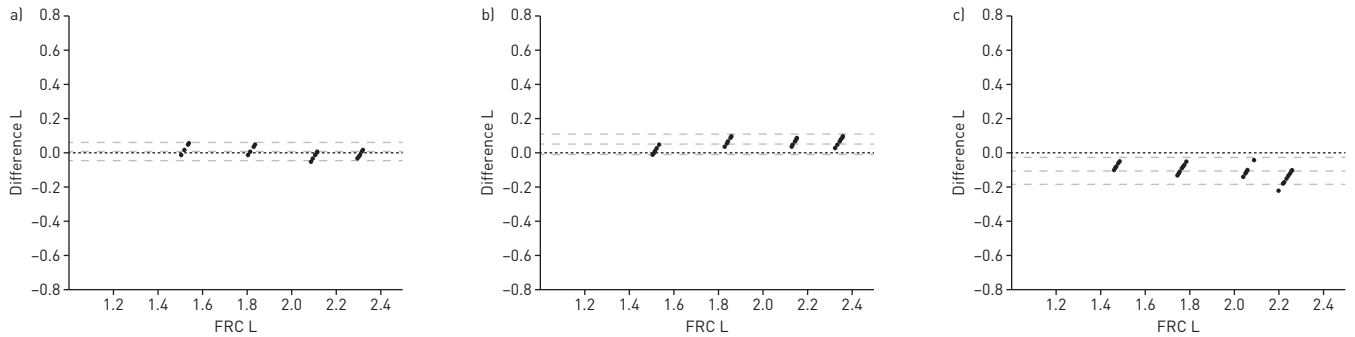


FIGURE 2 Bland–Altman plots of *in vitro* functional residual capacity (FRC) measurements on each device. Data are plotted as measured FRC minus lung model FRC, expressed as the absolute difference *versus* mean of measured and lung model FRC values. Absolute differences (circles), mean difference and upper and lower limits of agreement (mean difference \pm SD of differences) are shown as dashed lines. a) Woolcock in-house device; b) Exhalizer D device; c) EasyOne Pro LAB device analysed in Version 2.01.00.09.

measured *in vitro* FRC closest to the known lung model volume. The mean overestimation of *in vitro* FRC and FRC_{pleth} was 3% by the EM device, in comparison to a mean 21% underestimation by the ndd device. However, on reanalysis using ndd v.2.01, underestimation was reduced to 5% (*in vitro*) and 11% (*in vivo*), respectively. There were statistically significant differences in FRC measurements between commercial MBNW devices, although this difference was relatively small between EM and ndd v.2.01. Furthermore, the pattern of differences (*i.e.* over-estimation or under-estimation of FRC) between devices was consistent using both the *in vitro* physical lung model and the *in vivo* measurements.

Comparison to other studies

The Consensus recommendations based on expert opinion stated that 95% of the *in vitro* FRC measurements should be within 5% of the target volume [4]. WIMR and EM devices fulfilled this criterion; however, the ndd device did not achieve 5% accuracy for any measurements (v.2.00). When measurements were reanalysed using ndd v.2.01, 36% fulfilled this criterion. Both commercial devices were reported previously to be highly accurate in measuring FRC using a physical lung model [14]. This lung model was water-based and incorporated BTPS (body temperature, ambient pressure, saturated) correction, whereas ours did not. Despite this, our results showed the same pattern of over/underestimation *in vitro* and *in vivo*. This is consistent with their *in vivo* results in 10 healthy adults, where EM overestimated FRC_{pleth} by 14% and ndd underestimated by 23%. A recent paediatric study with healthy control and cystic fibrosis patients also reported the same pattern, *i.e.* higher FRC measurements using the EM device compared to the ndd device [15]. Other *in vitro* and *in vivo* paediatric studies exist [9, 11], but they involve different tracer gases, which limits comparison.

Impact of N₂ estimation method

Differences in FRC across the three devices are probably attributable to device and software variations. This includes the method used to calculate N₂ concentration. The WIMR device measures N₂ directly, whereas the EM device uses Dalton's law to compute N₂ by simple subtraction of other constituent gases in the expired air. These two methods were found to be similar. In contrast, the ndd v.2.00 device uses the concept of a prototype expirogram, derived from the shape of the molar mass versus expired volume curve in the early breaths of the washout. The expired N₂ volume for each breath is then determined by scaling

TABLE 2 Differences in *in vitro* functional residual capacity (FRC) measurements from the lung model functional residual capacity for four different lung volumes

Lung model FRC volume	Subjects	WIMR		EM		ndd v.2.00		ndd v.2.01	
		Absolute	Relative	Absolute	Relative	Absolute	Relative	Absolute	Relative
2.31	9	-0.01 \pm 0.02	-0.2 \pm 0.8%	0.07 \pm 0.02	3.3 \pm 1.0%	-0.54 \pm 0.14	-23.4 \pm 6.1%	-0.15 \pm 0.04	-6.3 \pm 1.7%
2.11	9	-0.01 \pm 0.02	-0.43 \pm 0.9%	0.06 \pm 0.02	2.9 \pm 0.9%	-0.39 \pm 0.12	-18.5 \pm 5.7%	-0.11 \pm 0.03	-5.2 \pm 1.5%
1.81	9	0.03 \pm 0.02	1.6 \pm 1.4%	0.07 \pm 0.02	3.7 \pm 1.0%	-0.44 \pm 0.11	-24.1 \pm 6.4%	-0.09 \pm 0.03	-4.7 \pm 1.6%
1.51	9	0.02 \pm 0.03	1.7 \pm 13.1%	0.02 \pm 0.02	1.1 \pm 1.3%	-0.26 \pm 0.05	-17.1 \pm 3.0%	-0.08 \pm 0.02	-5.0 \pm 1.2%

Data presented as mean \pm SD, unless otherwise stated. Absolute measurements are in litres. WIMR: Woolcock in-house device; EM: Exhalizer D device; ndd v.2.00: EasyOne Pro LAB device Version 2.00.01.05; ndd v.2.01: EasyOne Pro LAB device analysed in Version 2.01.00.09.

TABLE 3 Subject demographics and standard lung function measurements

	All subjects	Controls	Asthmatics
Subjects n	29	20	9
Males/females n	12/17	8/12	4/5
Age years	37±12	36±12	37±13
Height cm	172±12	174±12.5	168±10
BMI kg·m⁻²	24.1±3.63	23.9±3.48	24.6±4.10
Smoking history pack-years	2.5±3.2	2.75±3.31	0.70±0.00
FEV₁ L	3.66±0.93	3.75±0.93	3.47±0.95
FVC L	4.75±1.17	4.77±1.16	4.71±1.25
FEV₁/FVC	0.77±0.06	0.79±0.05	0.74±0.08
TLC L	6.51±1.38	6.55±1.41	6.42±1.41
FRC_{pleth} L	3.44±0.77	3.49±0.78	3.34±0.80

Data presented as mean±SD unless otherwise stated. BMI: body mass index; FEV₁: forced expiratory volume in 1 s; FVC: forced vital capacity; TLC: total lung capacity; FRC_{pleth}: functional residual capacity measured using body plethysmography.

the prototype expirogram to match the end-expiratory N₂ concentration for that breath. Potential inaccuracies could occur if the expirogram shape changed greatly during the course of the washout, as is typically observed in more severe obstructive airways disease. However, ndd v.2.01 uses a combination of molar mass measurements and Dalton's law to compute N₂, on a point-by-point basis for the entire expirogram. This could partly explain the improved accuracy of ndd v.2.01 and the reduced FRC discrepancy between the new software and the WIMR and EM devices.

Impact of other software changes

The other major change in ndd v.2.01 involve modifications to the estimation of delay between the flow and respective gas measurement points, which are more robust to variation in breathing patterns, particularly very brief pauses in flow (personal communication with the manufacturer). Differences in delay time potentially have a large effect on FRC calculations [20, 21].

Newer software versions were available at the time of writing (v.2.2.0.15 onwards). These involve major changes to the user interface; however, they fundamentally employ the same indirect N₂-based FRC calculation and delay estimation method. Our results thus have significant implications on the reanalysis of old data and are relevant to ndd v.2.01 as well as any newer versions. Furthermore, they illustrate the importance of analysis methods and any preprocessing algorithms that may be used, as well as the need for software transparency [22]. They also highlight the importance of ongoing validation and that standardisation efforts are gradually working.

It should be noted that this study was unable to confirm whether collecting data using ndd v.2.01 would have further improved the FRC accuracy for the ndd device. A large proportion of the *in vivo* tests reanalysed using the new software failed to meet end-of-test criteria, possibly due to the underestimation

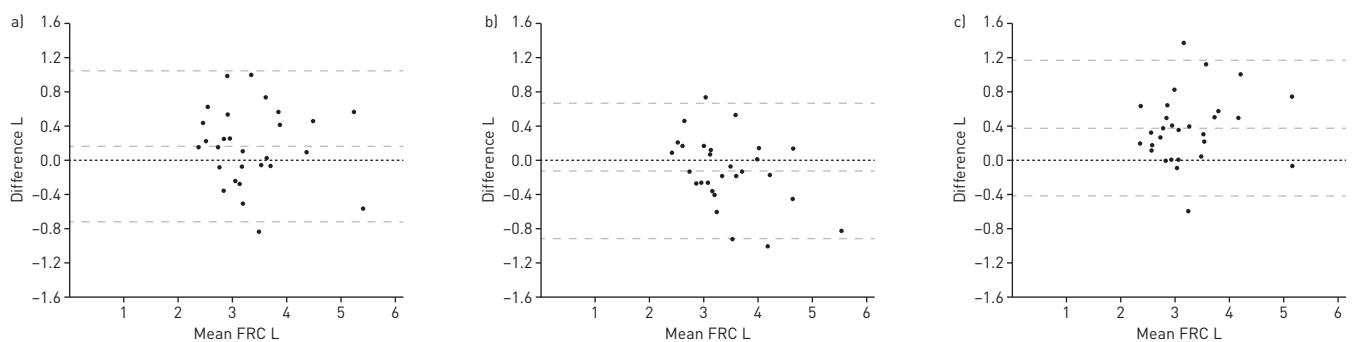


FIGURE 3 Bland-Altman plots of *in vivo* functional residual capacity (FRC) measurements on each device. Data are plotted as body plethysmography FRC (FRC_{pleth}) minus multiple-breath nitrogen washout (MBNW) FRC, expressed as the absolute difference versus mean of FRC_{pleth} and MBNW FRC. Absolute differences (circles), mean difference and upper and lower limits of agreement (mean difference±SD of differences) are shown as dashed lines. a) Woolcock in-house device; b) Exhalizer D device; c) EasyOne Pro LAB device analysed in Version 2.01.00.09.

TABLE 4 Differences in *in vivo* functional residual capacity (FRC) measurements from body plethysmography FRC on each device

	Subjects n	WIMR		EM		nnd v2.00		nnd v2.01	
		Absolute	Relative	Absolute	Relative	Absolute	Relative	Absolute	Relative
Controls	20	-0.13±0.46 [#]	-3.3±13.0%	0.17±0.41	4.7±11.8%	-0.73±0.36	-20.6±8.9%	-0.36±0.32	-10.1±8.4%
Asthmatics	9	-0.24±0.45	-6.8±13.8%	0.01±0.39	0.12±12.0%	-0.73±0.59	-20.7±15.3%	-0.42±0.57	-11.2±15.8%
All subjects	29	-0.17±0.45	-4.4±13.1%	0.12±0.40	3.3±11.8%	-0.73±0.43	-20.6±11%	-0.38±0.40	-10.5±10.9%

Data presented as mean±SD, unless otherwise stated. Absolute measurements are in litres. WIMR: Woolcock in-house device; EM: Exhalizer D device; nnd v.2.00: EasyOne Pro LAB device Version 2.00.01.05; nnd v.2.01: EasyOne Pro LAB device analysed in Version 2.01.00.09. [#]: missing data for one subject.

of N₂ in the old software. One could speculate that if acceptable end-tidal N₂ concentration had been met, *in vitro* FRC may have been slightly closer to the true lung volume and *in vivo* FRC closer to FRC_{pleth}.

Impact of patient factors

Other factors that may affect FRC measurements include the effect of breathing patterns during the washout. Different mouthpieces, resistances, the nature of the real-time displays of volume and patient breathing incentives, and even different open bypass systems delivering oxygen [15], may have affected breathing patterns. However, dead space correction was applied in each device and the same 1 L breathing protocol was used. More importantly, the direction of over/underestimation was preserved regardless of *in vitro* and *in vivo* FRC comparisons, *i.e.* independent of BTPS correction, breathing pattern and the presence of CO₂. Thus, it is unlikely that patient factors contributed significantly to the differences between devices. In addition, the differences between MBNW FRC and FRC_{pleth} on each device were of similar magnitude between healthy controls and subjects with asthma with well-preserved spirometry. It is unknown whether the same results would hold in subjects with more airway obstruction, or other patient groups such as chronic obstructive pulmonary disease (COPD).

Limitations

Our study had a number of limitations. First, we compared differences only in FRC between devices; however, LCI, S_{acin} and S_{cond} are the more clinically relevant MBNW indices. There are no known lung models to evaluate these other indices and a formal comparison of these indices *in vivo* was not performed. Nevertheless, high-quality FRC measurement is necessary for evaluation of these indices. Second, between-session variability for FRC has not been defined, so it is not known whether differences between *in vivo* measurements are within the limits of normal test variability. Third, the applicability of our results to the paediatric population is unknown. However, the simplistic design of the lung model could be easily adapted to a smaller syringe. Also, we only evaluated the standardised 1 L breathing protocol [23], so the effect of the free breathing protocol used by other groups [24], and especially in infants and children, is unknown. Furthermore, the syringe is an ATPD model rather than a BTPS model, which is less representative of the actual physiological situation but more practical for routine laboratory use. Despite this, the same pattern of accuracy was seen *in vitro* and *in vivo*, suggesting that incorporation of BTPS correction has a minimal effect on the relative errors reported. Finally, as discussed above, we did not evaluate the accuracy of data collected using the nnd v.2.01, which is scope for further work.

Conclusion

We have shown differences in the measurements of FRC between three MBNW devices in a physical syringe model of the lung. The syringe lung model used in this study was simple, portable and relatively easy to produce compared to that used in other studies [6, 25, 26]. It would allow a simple and practical way to calibrate or check the MBNW setup, because it tests the combined accuracy of volume and N₂ concentration measurement during a more realistic expirogram and the correct alignment of these signals. The use of such a syringe model may be beneficial for comparison between devices and laboratories and for quality assurance monitoring. FRC differences were also reflected *in vivo* in healthy and asthmatic subjects, in relation to plethysmographic FRC. While further work is required to improve accuracy, the *in vivo* differences observed are small and probably not clinically significant. Differences are likely to reflect the method of calculating N₂ concentration and other software factors. How these FRC errors translate to the accuracy of other indices has been explored for LCI [15], but not S_{acin} and S_{cond} so far. Nevertheless, our results show that the state of the art is closer to achieving better comparability and standardisation for FRC accuracy across existing MBNW devices.

Acknowledgements

The authors acknowledge the contribution of Gunnar Unger, Martin Turner, Aaron Skelsey and Sunny Ye, who were involved with the development of the syringe lung model. Both manufacturers for the Eco Medics and ndd devices were consulted for the technical accuracy of the manuscript. The authors acknowledge the contribution of Christian Buess, who provided additional technical information regarding the ndd software where indicated in the manuscript. Neither manufacturer influenced the study design, results or interpretation.

References

- 1 Verbanck S, Schuermans D, Meysman M, *et al.* Noninvasive assessment of airway alterations in smokers. *Am J Respir Crit Care Med* 2004; 170: 414–419.
- 2 Jetmalani K, Chapman DG, Thamrin C, *et al.* Bronchodilator responsiveness of peripheral airways in smokers with normal spirometry. *Respirology* 2016; 21: 1270–1276.
- 3 Lahzami S, Schoeffel RE, Pechey V, *et al.* Small airways function declines after allogeneic haematopoietic stem cell transplantation. *Eur Respir J* 2011; 38: 1180–1188.
- 4 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.
- 5 Pillow JJ, Ljungberg H, Hulskamp G, *et al.* Functional residual capacity measurements in healthy infants: ultrasonic flow meter versus a mass spectrometer. *Eur Respir J* 2004; 23: 763–768.
- 6 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.
- 7 Owens CM, Aurora P, Stanojevic S, *et al.* Lung Clearance Index and HRCT are complementary markers of lung abnormalities in young children with CF. *Thorax* 2011; 66: 481–488.
- 8 Fuchs SI, Buess C, Lum S, *et al.* Multiple breath washout with a sidestream ultrasonic flow sensor and mass spectrometry: a comparative study. *Pediatr Pulmonol* 2006; 41: 1218–1225.
- 9 Jensen R, Stanojevic S, Gibney K, *et al.* Multiple breath nitrogen washout: a feasible alternative to mass spectrometry. *PLoS ONE* 2013; 8: e56868.
- 10 Schibler A, Hall GL, Businger F, *et al.* Measurement of lung volume and ventilation distribution with an ultrasonic flow meter in healthy infants. *Eur Respir J* 2002; 20: 912–918.
- 11 Schmidt A, Yammine S, Proietti E, *et al.* Validation of multiple-breath washout equipment for infants and young children. *Pediatr Pulmonol* 2015; 50: 607–614.
- 12 Summermatter S, Singer F, Latzin P, *et al.* Impact of software settings on multiple-breath washout outcomes. *PLoS ONE* 2015; 10: e0132250.
- 13 Anagnostopoulou P, Yammine S, Schmidt A, *et al.* False normal Lung Clearance Index in infants with cystic fibrosis due to software algorithms. *Pediatr Pulmonol* 2015; 50: 970–977.
- 14 Raaijmakers L, Jensen R, Stanojevic S, *et al.* *In vitro* and *in vivo* comparison of two multiple breath nitrogen washout devices. *Eur Respir J* 2015; 46: Suppl. 59, PA1262.
- 15 Poncin W, Singer F, Aubriot A-S, *et al.* Agreement between multiple-breath nitrogen washout systems in children and adults. *J Cyst Fibros* 2016; 16: 258–266.
- 16 Verbanck S, Schuermans D, Van Muylem A, *et al.* Ventilation distribution during histamine provocation. *J Appl Physiol* 1997; 83: 1907–1916.
- 17 Downie SR, Salome CM, Verbanck S, *et al.* Ventilation heterogeneity is a major determinant of airway hyperresponsiveness in asthma, independent of airway inflammation. *Thorax* 2007; 62: 684–689.
- 18 Hardaker KM, Downie SR, Kermod JA, *et al.* Predictors of airway hyperresponsiveness differ between old and young patients with asthma. *Chest* 2011; 139: 1395–1401.
- 19 Wanger J, Clausen JL, Coates A, *et al.* Standardisation of the measurement of lung volumes. *Eur Respir J* 2005; 26: 511–522.
- 20 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.
- 21 Egger B, Jost K, Anagnostopoulou P, *et al.* Lung clearance index and moment ratios at different cut-off values in infant multiple-breath washout measurements. *Pediatr Pulmonol* 2016; 51: 1373–1381.
- 22 Mahar RK, Vukcevic D, King L, *et al.* Lack of transparency in software used to analyze multiple breath washout data. *Pediatr Pulmonol* 2016; 51: 1108–1110.
- 23 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.
- 24 Kjellberg S, Houltz BK, Zetterström O, *et al.* Clinical characteristics of adult asthma associated with small airway dysfunction. *Respir Med* 2016; 117: 92–102.
- 25 Gonem S, Singer F, Corkill S, *et al.* Validation of a photoacoustic gas analyser for the measurement of functional residual capacity using multiple-breath inert gas washout. *Respiration* 2014; 87: 462–468.
- 26 Williams EM, Gale LB, Oakley PA, *et al.* Development of a concentric water-displacement model lung. *J Biomed Eng* 1993; 15: 420–424.