ERJ open research
THE BEST IN OPEN ACCESS BASIC, TRANSLATIONAL & CLINICAL RESPIRATORY RESEARCH

Early View

Original article

# Controlled *versus* free breathing for multiple breath nitrogen washout in healthy adults

Blake M. Handley, Edward Jeagal, Robin E. Schoeffel, Tanya Badal, David G. Chapman, Catherine E. Farrow, Gregory G. King, Paul D. Robinson, Stephen Milne, Cindy Thamrin

Please cite this article as: Handley BM, Jeagal E, Schoeffel RE, *et al*. Controlled *versus* free breathing for multiple breath nitrogen washout in healthy adults. *ERJ Open Res* 2020; in press (https://doi.org/10.1183/23120541.00435-2020).

**MANUSCRIPT TITLE:**

**Controlled vs free breathing for multiple breath nitrogen washout in healthy adults**

**AUTHORS:**

**Blake M HANDLEY[1,2], Edward JEAGAL[2], Robin E SCHOEFFEL[1], Tanya BADAL[2], David G CHAPMAN[2,3], Catherine E FARROW[2,6], Gregory G KING[1,2,4], Paul D ROBINSON[5], Stephen MILNE[4,7]*, Cindy THAMRIN[2,4]***

1. Department of Respiratory Medicine, Royal North Shore Hospital, St Leonards, NSW, Australia
2. Woolcock Institute of Medical Research, University of Sydney, Glebe NSW, Australia
3. Translational Airways Group, School of Life Sciences, Faculty of Science, University of Technology Sydney, Ultimo NSW, Australia
4. Faculty of Medicine and Health, University of Sydney, Camperdown NSW, Australia
5. Department of Respiratory Medicine, The Children's Hospital at Westmead, Westmead NSW, Australia.
6. Department of Respiratory and Sleep Medicine, Westmead Hospital, Westmead NSW, Australia
7. Centre for Heart Lung Innovation, St Paul's Hospital and Division of Respiratory Medicine, University of British Columbia, Vancouver BC, Canada

***joint Senior Authors with equal contributions**

**CORRESPONDING AUTHOR:**

Dr Cindy Thamrin

Airways Physiology and Imaging Group

Woolcock Institute of Medical Research

431 Glebe Point Rd

Glebe NSW 2037

AUSTRALIA

E: cindy.thamrin@sydney.edu.au

**TAKE-HOME MESSAGE:**

Indices derived from 'free breathing' and 'controlled breathing' MBNW protocols are not necessarily comparable, and differences may be related to breathing patterns. These findings have implications for the ongoing clinical implementation of MBNW.

**KEYWORDS:**

Multiple breath nitrogen washout; ventilation heterogeneity

**WORD COUNT:** 3,801 words

**ABSTRACT**

Multiple breath nitrogen washout (MBNW) quantifies ventilation heterogeneity. Two distinct protocols are currently used for MBNW testing: 'controlled breathing', with targeted tidal volume (VT) and respiratory rate (RR); and 'free breathing', with no constraints on breathing pattern. Indices derived from the two protocols (functional residual capacity [FRC], lung clearance index [LCI], Scond, Sacin) have not been directly compared in adults. We aimed to determine whether MBNW indices are comparable between protocols, to identify factors underlying any between-protocol differences, and to determine the between-session variabilities of each protocol.

We performed MBNW testing by both protocols in 27 healthy adult volunteers, applying the currently-proposed correction for VT to Scond and Sacin derived from free breathing. To establish between-session variability, we repeated testing in 15 volunteers within 3 months. While FRC was comparable between controlled vs free breathing (3.17(0.98) vs 3.18(0.94) L,p=0.88), indices of ventilation heterogeneity derived from the two protocols were not, with poor correlation for Scond (r=0.18,p=0.36) and significant bias for Sacin (0.057(0.021)$L^{-1}$ vs 0.085(0.038)$L^{-1}$,p=0.0004). Between-protocol differences in Sacin were related to differences in the breathing pattern, i.e. VT (p=0.004) and RR (p=0.01), rather than FRC. FRC and LCI showed good between-session repeatability, but Scond and Sacin from free breathing showed poor repeatability with wide limits of agreement.

These findings have implications for the ongoing clinical implementation of MBNW, as they demonstrate that Scond and Sacin from free breathing, despite VT correction, are not equivalent to the controlled breathing protocol. The poor between-session repeatability of Scond during free breathing may limit its clinical utility.

## INTRODUCTION

Increased ventilation heterogeneity is a characteristic physiological abnormality in respiratory diseases such as asthma and COPD [1, 2]. The multiple breath nitrogen washout (MBNW) test is an increasingly-available method of quantifying ventilation heterogeneity. MBNW is conducted by breathing 100% oxygen ($O_2$) which then 'washes out' the resident nitrogen ($N_2$) in the lung. Analysis of the exhaled $N_2$ concentration versus exhaled volume of each breath then allows calculation of a number of parameters: a global measure of heterogeneity known as the lung clearance index (LCI); heterogeneity arising predominantly within the conducting airways (Scond); heterogeneity arising predominantly in the more distal/intra-acinar airways (Sacin); and functional residual capacity (FRC).

Two distinct breathing protocols are currently used for MBNW testing. One is a 'controlled breathing' protocol whereby the patient maintains a consistent tidal volume (VT) (approximately 1 L) and respiratory rate (RR) (8-12 breaths/min) [3] through the use of a visual incentive screen and real-time feedback from the test operator. These breathing constraints are applied to all individuals equally, regardless of lung size; consequently, a given patient's breathing pattern during the test may deviate significantly from their usual resting breathing pattern. The alternative is a 'free breathing' protocol, with no constraints on VT or RR, i.e. the individual determines their own breathing depth and frequency. While this makes the method more suited to paediatric testing [3], variability in breathing patterns within or between individuals may have an impact on indices of ventilation heterogeneity [4-6] and their repeatability. A correction for lung size and breath-to-breath variation in VT has been proposed for this method [7], but this has not been formally validated against a controlled breathing approach. Standard reference equations are available for both methods to guide interpretation of MBNW indices, which account for factors such as age, height and sex [8-11], but little is known about their comparability.

Differences in how controlled and free breathing MBNW protocols are conducted and analysed may therefore limit the comparability of indices derived from them. A direct

comparison of the protocols in children [12] showed significant differences in LCI and Scond; the authors concluded that this was due to a reduction in end-expiratory lung volume (EELV) induced by controlled breathing. However, the results suggested that other patient-related factors (e.g. lung size, body weight) may contribute to the discrepancies. Indices derived from the two protocols have not been directly compared in adults, nor have their respective between-session variabilities.

Therefore, in a sample of healthy adults, we aimed to determine: 1) whether controlled and free breathing MBNW protocols provide equivalent FRC and indices of ventilation heterogeneity; 2) the influence of patient-related factors (anthropometrics and/or breathing pattern) on any observable differences; and 3) if the choice of protocol influences between-session variability, which would allow us to interpret any between-protocol differences. Our overall hypothesis was that controlled and free breathing MBNW would produce comparable measurements of ventilation heterogeneity, with the secondary hypothesis that any differences would be, in part, due to patient-related factors.

## METHODS

*Study overview*

We recruited and studied volunteers aged ≥18 years in the respiratory function laboratories at two sites (RNSH and WIMR, Sydney, Australia; see Supplement for inclusion/exclusion criteria) from a convenience sample of volunteers (predominantly hospital staff) between April 2018 and November 2019. Our target sample size was n=25, comparable to the control group of a previous study [12]. The RNSH Human Research Ethics Committee approved the study (LNR/16/HAWKE/11). After obtaining written informed consent to be tested and for their data to be used for research purposes, participants first underwent standard (pre-bronchodilator) lung function testing to confirm they had normal lung function. After a period of rest of at least 10 minutes, participants underwent MBNW testing by both breathing protocols conducted in a randomised order. Each participant was invited to return for repeat

testing, at their convenience, within the next three months. This follow-up testing was performed in an identical manner, including the order of MBNW protocol testing, for consistency.

*Standard lung function testing*

We measured spirometry and lung volumes according to ATS/ERS quality criteria using a Jaeger MasterScreen PFT (Vyaire Medical V2.21.4) device, with comparison to reference values [13, 14] for plethysmography.

*MBNW testing*

Detailed descriptions of the device and testing procedure are found in the Supplement. After at least 10 minutes of rest for the participant, we conducted MBNW tests using the Exhalyzer D with Spiroware v3.1.6 (Eco Medics AG, Duernten, Switzerland).

Participants performed both controlled and free breathing protocols in succession, in a random order (as determined by a computer-based random number generator), according to current international consensus recommendations [3]. All trials started with a period of normal relaxed breathing on room air in order to establish a stable EELV. For the controlled breathing protocol, the operator then instructed the participant to breathe at a VT of 0.95-1.3 L and at RR 8-12/min with the use of the visual incentive screen within Spiroware (described by Verbanck [15]); once the operator was satisfied that a stable breathing pattern and EELV had been achieved, they commenced the washout by switching the circuit to 100% $O_2$ and the participant maintained the same breathing pattern for the duration of the washout phase. For the free breathing protocol, the operator instructed the participant to “continue to breathe in a normal relaxed manner” through both the pre-phase and washout-phase, with the visual incentive screen switched off and without any additional coaching during the trial. Participants repeated the test until at least 3 technically acceptable trials with FRC values within ±10% of the mean were achieved for each protocol.

*MBNW analysis*

Full details are provided in the Supplement. We analysed MBNW data using Spiroware software (v.3.1.6). While the test operator performed preliminary analysis on individual trials during the testing session, a single investigator re-analysed the data *post hoc* for all participants as a batch in order to ensure a consistent approach to analysis.

To differentiate between measurements made by the different protocols, we have subscripted all indices with CB or FB to indicate controlled and free breathing, respectively. Under the free breathing protocol, Scond and Sacin are adjusted for VT and are thus denoted in the literature and Spiroware software as Scond*VT and Sacin*VT; however, we refer to them here simply as $Scond_{FB}$ and $Sacin_{FB}$, respectively.

*Statistical analyses*

We compared FRC, LCI, Scond, and Sacin measured by the two breathing protocols using paired t tests and Pearson correlation. Additionally, we compared FRC from either protocol against the gold-standard $FRC_{pleth}$. To investigate for bias, we generated Bland-Altman plots as the between-protocol difference (free breathing minus controlled breathing) versus the average, plotting the mean difference and 95% limits of agreement (95% LOA). We then performed linear regression of the difference versus average to determine any proportional bias. To examine the effects of various predictors (age, sex, height, BMI, mean RR from free breathing, mean VT from free breathing, and $FRC_{pleth}$) on between-protocol difference, we performed linear regression of the difference versus each predictor. To determine within-session variability, we calculated the coefficient of variation (CoV) from the three washout trials. To determine between-session variability, we calculated the difference (Visit 2 minus Visit 1) and 95% LOA separately for each protocol. We also report the between-session intra-class correlation coefficients (ICC), calculated using a two-way mixed effects ANOVA model based on absolute agreement, multiple measurements (k=3) [16], We set statistical

significance at $p<0.05$. Results are presented as mean(standard deviation [SD]) unless otherwise stated.

## RESULTS

*Participant characteristics and breathing patterns*

We studied 27 non-smoking, healthy volunteers (22 at the RNSH site, 5 at the WIMR site) with a median age of 34 (range 19-65) years and spirometry/lung volumes within the limits of normal (Table 1). Since MBNW parameters are likely to be device- and protocol-specific [17, 18], we were unable to compare all measurements to a single set of reference equations. However, MBNW indices from the free breathing protocol were within the ranges of normal derived from the same testing device (Table S1) [9].

Compared to controlled breathing, free breathing produced smaller mean VT (mean(SD) difference -0.24(0.33) L, p=0.0006), faster mean RR (1.6(3.4) breaths/min, p=0.02), smaller minute ventilation (-2.4(3.1) L/min, p<0.0001), greater cumulative expired volume (CEV) (1.84 L, p=0.006), but similar washout times (p=0.17). Four out of the 27 participants required coaching prior to commencing the free breathing protocol due to inadequate phase III. The average proportions of washout breaths excluded from analysis due to inadequate phase III were 2.7% for controlled breathing and 3.9% for free breathing.

*FRC was comparable between protocols, and with plethysmography*

There was no significant difference in mean FRC measured by either MBNW protocol ($FRC_{CB}$ 3.17(0.98) vs $FRC_{FB}$ 3.18(0.94), p=0.88). $FRC_{CB}$ and $FRC_{FB}$ were strongly correlated (r=0.94, p<0.0001; Figure 1A). There was no evidence of bias in the absolute (Figure 1B) or percentage difference (Figure S1).

There were no significant differences between mean $FRC_{CB}$ or $FRC_{FB}$ and mean $FRC_{pleth}$ (p=0.83 and p=0.86, respectively). Both $FRC_{CB}$ and $FRC_{FB}$ were strongly correlated with

$FRC_{pleth}$ (r=0.84 and r=0.92, respectively, p<0.0001 for both; Figure 2). Within-subject differences in FRC between the two MBNW protocols were not related to the individual's $FRC_{pleth}$ (p=0.55, Figure S2-A). Similarly, within-subject differences in FRC between the protocols were not related to age, sex, height, mean VT or mean RR from free breathing, but there was a trend towards an effect of BMI (p=0.07, Figure S2-B).

*LCI was comparable between protocols*

Mean LCI was significantly lower with controlled breathing ($LCI_{CB}$ 7.2(0.58) vs $LCI_{FB}$ 7.55(0.81), p=0.0004), however the mean difference was small. There was a strong correlation between the protocols (r=0.84, p<0.0001; Figure 3-A). However, there was evidence of proportional bias (between-protocol difference increased with LCI, p=0.004; Figure 3-B).

*Scond was poorly correlated between protocols*

Overall, mean Scond measured by both protocols was not significantly different ($Scond_{CB}$ 0.017(0.009) vs $Scond_{FB}$ 0.018(0.01) $L^{-1}$, p=0.74). However, there was no significant correlation between the protocols (r=0.18, p=0.36, Figure 4-A). The Bland-Altman plot revealed large variance in between-protocol differences and possible (but non-significant) proportional bias (p=0.45, Figure 4-B).

*Sacin showed significant between-protocol differences*

Mean Sacin was significantly lower with controlled breathing ($Sacin_{CB}$ 0.057(0.021) $L^{-1}$ vs $Sacin_{FB}$ 0.085(0.038) $L^{-1}$, p=0.0004). The correlation between the protocols was borderline significant (r=0.37, p=0.06; Figure 5-A). The Bland-Altman plot revealed significant proportional bias (between-protocol difference increased with Sacin, p=0.002; Figure 5-B). Linear regression showed that within-subject differences in Sacin between the two protocols were related to the breathing pattern. Specifically, the greater the deviation in mean VT or mean RR between the protocols, the larger the discrepancy in Sacin (p=0.004 and p=0.01,

respectively; Figures 5-C and 5-D), such that participants who breathed shallower or faster during the free breathing protocol had greater apparent $Sacin_{FB}$.

*Breathing protocol influenced between-session variability of MBNW indices*

15 volunteers underwent repeat testing; the baseline (Visit 1) characteristics of this subgroup were similar to those of the group as a whole, except for a narrower age range (Table 1). The mean(SD) time between sessions was 5.9(3.3) weeks. Within-session and between-session variability measures for both protocols are presented in Table 2. There were no differences seen in within-session CoV between protocols in FRC (p=0.677) or LCI (p=0.157). In terms of between-session variability, the free breathing protocol showed relatively greater variability in LCI, Scond and Sacin, as indicated by numerically greater mean differences and wider 95% LOAs, with very poor between-session ICC seen in Scond and Sacin. The controlled breathing protocol showed wider 95% LOA for FRC. Bland-Altman plots did not suggest any proportional bias for either protocol (Figures S3, S4).

## DISCUSSION

*Summary of results*

In this study in healthy adults comparing two commonly-used MBNW breathing protocols, we found that: 1) FRC was comparable between the two protocols, however indices of ventilation heterogeneity based on phase III slopes were not – Scond was poorly correlated between the two protocols and Sacin were systematically higher under the free breathing protocol, whereas LCI was only marginally higher; 2) there was greater between-protocol discrepancy in Sacin in subjects whose free breathing pattern deviated from that of the volume- and frequency-controlled protocol; and 3) Scond and Sacin exhibited poorer between-session repeatability under the free breathing protocol.

*Lack of discrepancy in FRC*

Yammine et al [12] demonstrated in children that Scond and LCI obtained from a controlled breathing protocol were higher than when obtained by free breathing, and that the discrepancy was related to a decrease in FRC during controlled breathing. Thus, the authors argued that a controlled breathing protocol may overestimate heterogeneity by inadvertently causing individuals to change the EELV at which they would otherwise normally breathe. In children, a 1 L VT represents a significant proportion of TLC; the high VT may increase ventilation heterogeneity due to the recruitment of normally non-ventilated lung units [19], or due to closure of lung units in the dependent regions if the child expires to below their normal end-expiratory lung volume [20], or some combination of both.

Our current findings in healthy adults do not support the argument that FRC is altered with controlled breathing, as we showed that FRC estimated by the two protocols was in fact comparable, with no systematic bias. Furthermore, both were comparable to the 'gold standard', i.e. $FRC_{pleth}$. A number of factors may explain the differences between our findings and those of Yammine et al. For example, the effect of variation in FRC may be relatively mild in the adults in our study compared to that in children, and thus insufficient to cause significant differences in heterogeneity. We observed that the variability in between-protocol differences (assessed by the LOA) in FRC were comparable to the between-session variabilities of either protocol, though interestingly between-session LOA was numerically higher under the controlled breathing protocol. Alternatively, our standard procedure for the controlled breathing protocol includes an initial period of unconstrained breathing to allow the individual's end-expiratory lung volume to stabilise before commencing the washout. This is in line with current ATS/ERS recommendations [3] and may have ensured FRC differences were kept minimal between protocols.

*Discrepancy in ventilation heterogeneity indices*

While LCI – an index of global heterogeneity that is independent of SIII – was comparable between protocols (with a small bias), we found that SIII-dependent ventilation heterogeneity partitioned into proximal/conducting (Scond) and distal/acinar (Sacin) zones was not. Scond was poorly correlated between protocols, perhaps driven by a large between-session variability. Sacin was significantly different between protocols; moreover, Sacin obtained by free breathing was higher than that obtained by controlled breathing in individuals with relatively small VT or high RR. Between-protocol differences in Sacin appear to be independent of age, height or lung size, i.e. a change in breathing pattern was the dominant driver of the discrepancy.

There are two possible explanations for this. First, the effect could be purely methodological: when breaths are too 'shallow', estimation of SIII is rendered invalid as the $N_2$ expirogram either never reaches a plateau or attains a plateau that is too short for reliable curve fitting, which would subsequently affect the derivation of Sacin and Scond. Anectodal evidence suggests this occurs frequently. Methods to more reliably estimate SIII exist [21], but are still ultimately dependent on the presence of distinct phase II and III regions in the expirogram and an adequate portion of phase III. More useful may be methods of quantifying ventilation heterogeneity that do not rely on SIII slope estimation [22, 23] – these may help not just in free breathing but also in pathologies where SIII is often difficult to define. The feasibility and clinical relevance of these methods should be a direction for future investigation.

Alternatively, the effect may be physiological: when patients breathe shallower and faster, the fast-emptying lung compartments will increase their $N_2$ washout while the slow-emptying lung compartments will be unable to empty completely, thus increasing apparent heterogeneity estimated by Sacin and Scond. Indeed, the foundational studies on ventilation distribution clearly demonstrated an effect of breathing pattern (including VT) on MBNW indices [4]. Recently, Ratjen and colleagues [24] investigated the effects of altered VT on MBNW indices and found that shallow breathing (compared to 'unrestricted' breathing) significantly increased LCI. The authors speculated that this was due to the effects of

increased dead space to VT ratio on the determinants of LCI (both CEV and FRC). Notably, these investigators did not examine the impact on Scond or Sacin.

We found that on average, VT was shallower, RR was faster and CEV higher in the free breathing protocol in our study participants. The change in CEV would be expected, since smaller breaths may reduce alveolar gas mixing efficiency and therefore increase the volume (and time, though this was not significant) required to reach end-of-test criteria. This likely drove the changes in LCI we observed – indeed, differences in LCI were directly proportional to differences in CEV (Figure S5). However, it should be noted that the differences seen in LCI were very small relative to its magnitude, and unlikely to be clinically significant. Furthermore, we found a weak relationship between differences in Scond and differences in CEV; this was not seen with Sacin (Figure S6). However, when plotted against differences in CEV/FRC (i.e. LCI, which corrects CEV for lung volume), the relationship with Sacin became significant. This supports the interpretation that the differences seen in Sacin were driven by changes in ventilation heterogeneity, rather than alterations to lung volume.

*Variability of MBNW indices*

We also report between-session variability over 2-10 weeks for both protocols, using the same commercially-available device. This allowed us to compare between-protocol differences in MBNW indices against their short-term variability. Within-session CoV of FRC and LCI was similar to published values for the free-breathing protocol [25], using the same device. However the between-session coefficient of repeatability (equal to 1.96xSD of the mean differences) of LCI, Scond, and Sacin were greater in our study, which may reflect the different time between sessions (weeks/months vs days). There is a paucity of repeatability data in health, and previous reports may not be generalizable due to persistent between-device differences [17, 18, 26]. Nevertheless, the excellent between-session repeatability for FRC and LCI [25, 27-29] but poor repeatability in Sacin and Scond [25, 28] has been noted with the free breathing protocol by other investigators, regardless of device, tracer gas, or

disease. The higher between-session repeatability values seen in Scond and Sacin with the controlled breathing protocol are consistent with other studies [30, 31].

It is also worth noting that FRC and LCI are derived from successive cumulation of volume at each breath, i.e. they are integrated measures, which are more robust to noise. On the other hand, Scond and Sacin are estimated from slopes, i.e. differentiated measures, which are inherently susceptible to noise. Furthermore, unlike FRC and LCI, the values of Scond and Sacin are very close to zero, which may explain the very low between-session ICC values observed. These may be fundamental reasons for the greater variability seen in Scond and Sacin but not FRC and LCI, and this variability becomes even more pronounced when variability in tidal breathing is introduced. This higher variability may also drive the differences seen between the two protocols. We also note the potential for these differences to be further exaggerated in disease – this and the high between-session variability observed may limit the clinical utility of SIII indices derived from free breathing. These speculations need to be confirmed in further investigations.

*Significance*

Both MBNW breathing protocols are designed to measure ventilation heterogeneity during resting tidal breathing. The controlled breathing protocol was originally devised based on modelling studies in which convection-dependent and diffusion-convection interaction-dependent mechanisms of airflow in the lung (the basis of Scond and Sacin, respectively) were first described [32]. The controlled breathing protocol standardises the VT at which ventilation distribution is assessed, and also ensures that expiration occurs well into the alveolar plateau so that a reliable SIII estimation can be made. This standardisation potentially minimises variability between tests. However, this ‘artificial’ pattern of breathing may distribute the ventilation differently in lungs of different sizes i.e. in shorter versus taller people. These differences may then affect the measurements of MBNW indices such that they no longer reflect true resting or ‘natural’ ventilation distribution. In contrast, under a free breathing protocol, breathing occurs at (or closer to) natural resting breathing pattern; thus,

the MBNW indices obtained arguably better reflect the indvidual's native ventilation distribution during normal tidal breathing.

Our results lend weight to the idea that controlled breathing imposes a condition that is different to the individual's natural resting breathing pattern, and that MBNW indices derived from this protocol may not necessarily reflect their 'native' ventilation heterogeneity. Specifically, breathing at 1 L tended to improve heterogeneity measured by Sacin in individuals whose VT during free breathing was less than 1 L, or whose RR during free breathing was faster than the prescribed 8-12 per minute. So, which is the 'better' test? We suggest that it is whichever provides the highest sensitivity in detecting disease or assessing treatment. The 'better' test may even be disease-specific, and warrants further investigation.

Since these differences are apparent despite the use of VT-corrected versions of Scond and Sacin, our study suggests that the linear VT correction currently employed for the free breathing protocol may not be adequate. Again, methods which assess ventilation heterogeneity without the need for estimating SIII, ideally in a manner that is independent of VT [22], may be useful here.

*Limitations*

Our study has several limitations to be taken into consideration. First, a number of the volunteers could be considered 'trained' in lung function testing. However, it is unlikely that prior experience would influence the largely effort-independent MBNW test. Second, 4/27 participants required coaching before or after free breathing testing due to insufficient expired volume for phase III slope estimation. Thus, there is a chance our results may actually underestimate the true discrepancy between the two protocols, particularly in shallow breathers [24] or in lung disease where increased dead space further complicates phase III slope estimation [33]. Third, our results only apply to a healthy population tested on a single MBNW device. There are likely to be differences between equipment [17, 18] and

between populations (e.g. children, older adults, lung disease) that limits the generalisability of our results.

*Conclusion*

We have demonstrated that while controlled and free breathing protocols are equivalent in terms of FRC and LCI, phase III-derived MBNW indices of ventilation heterogeneity are lower (i.e. better) when measured by the controlled breathing method. This effect seems to be driven largely by deviations in breathing pattern between the two protocols, particularly in patients who either breath faster or shallower than the criteria imposed by the controlled breathing protocol. Future work needs to be carried out in diseased patients to see if these observations hold true. Our study sheds light on potential physiological mechanisms behind these differences, and the overall interpretation of ventilation heterogeneity measured by MBNW. In better characterising the difference between the two protocols, these findings also help facilitate the ongoing efforts to standardise MBNW as an emerging clinical test.

**ACKNOWLEDGEMENTS**

We wish to acknowledge the help of the study participants, who kindly volunteered their valuable time for the study.

**FINANCIAL SUPPORT**

There were no direct financial sponsors for the submitted work. SM is supported by the MITACS Accelerate program.

**DISCLOSURES**

GGK reports grants, personal fees, non-financial support and other from AstraZeneca, Boehringer Ingelheim, CycloPharm, GlaxoSmithKline, Novartis, Menarini, and MundiPharma, and grants from National Health & Medical Research Council and the Asthma Foundation, outside the submitted work. SM reports personal fees from Novartis,

personal fees from Boehringer-Ingelheim, outside the submitted work. BMH, EJ, RES, DGC, CEF, PDR and CT have nothing to disclose.

## TABLES

**Table 1. Participant characteristics, lung function and MBNW parameters.** Values are mean(SD) unless otherwise stated. For the repeatability subgroup, results are from the first visit. Reference equations for predicted values from †Quanjer 2012 [13] and ‡Quanjer [14]. BMI, body mass index; $FEV_1$, forced expiratory volume in 1 s; FVC, forced vital capacity; TLC, total lung capacity; FRC, functional residual capacity; LCI, lung clearance index; Scond, conductive zone ventilation heterogeneity; Sacin, acinar zone ventilation heterogeneity; VT, mean tidal volume across a measurements; CEV, cumulative expired volume. Subscripts indicate the testing method/protocol: Pleth, body plethysmography; CB, controlled breathing protocol; FB, free breathing protocol; MBNW, multiple breath nitrogen washout.

| | **Whole group** | **Repeatability subgroup** |
|---|---|---|
| Females:Males, n | 11:16 | 6:9 |
| Age, yrs (range) | 34 (19-65) | 30 (23-41) |
| BMI, $kg/m^2$ | 24.6(3.4) | 25.1(4.2) |
| | | |
| **Lung function** | | |
| $FEV_1$, % predicted† | 105(14) | 101(30) |
| $FEV_1$/FVC, % | 83(6) | 84(5) |
| $TLC_{Pleth}$, % predicted‡ | 101(23) | 107(11) |
| $FRC_{Pleth}$, % predicted‡ | 97(27) | 104(20) |
| | | |
| **MBNW parameters** | | |
| *Controlled breathing* | | |
| $FRC_{CB}$, L | 3.17(0.98) | 3.23(0.97) |
| $LCI_{CB}$ | 7.2(0.58) | 7.13(0.5) |
| $Scond_{CB}$, $L^{-1}$ | 0.017(0.009) | 0.02(0.01) |
| $Sacin_{CB}$, $L^{-1}$ | 0.057(0.022) | 0.063(0.021) |
| VT, mL | 1124(37) | 1127(37) |
| $VT/FRC_{MBNW}$ | 0.40(0.15) | 0.39(0.14) |
| RR, breaths/minute | 12.21(1.36) | 11.9(1.53) |
| | | |
| *Free breathing* | | |
| $FRC_{FB}$, L | 3.18(0.94) | 3.25(0.96) |
| $LCI_{FB}$ | 7.55(0.81) | 7.4(0.79) |
| $Scond_{FB}$, $L^{-1}$ | 0.018(0.01) | 0.018(0.013) |
| $Sacin_{FB}$, $L^{-1}$ | 0.085(0.038) | 0.091(0.04) |
| VT, mL | 880(325) | 912(304) |
| $VT/FRC_{MBNW}$ | 0.29(0.11) | 0.3(0.13) |
| RR, breaths/minute | 13.78(3.43) | 13.18(3.46) |

**Table 2. Within- and between-session variability for the controlled and free breathing protocols.** Mean differences are Visit 2 minus Visit 1. CoV, coefficient of variation; 95% LOA, 95 % limits of agreement; ICC, intra-class correlation coefficient; FRC, functional residual capacity; LCI, lung clearance index; Scond, conducting airways ventilation heterogeneity; Sacin, distal/intra-acinar airways ventilation heterogeneity.

| | **Within-session CoV** | **Between-session mean difference(SD)** | **95% LOA** | **Between-session ICC** |
|---|---|---|---|---|
| | **Controlled Breathing** | | | |
| $FRC_{CB}$, L | 3.3(2.9) % | -0.03(0.52) | -1.04, 0.98 | 0.931 |
| $LCI_{CB}$ | 2.5(2.4) % | 0.06(0.47) | -0.86, 0.98 | 0.812 |
| $Scond_{CB}$, $L^{-1}$ | - | -0.001(0.011) | -0.017, 0.015 | 0.836 |
| $Sacin_{CB}$, $L^{-1}$ | - | -0.003(0.018) | -0.037, 0.031 | 0.835 |
| | **Free Breathing** | | | |
| $FRC_{FB}$, L | 3.6(2.3) % | -0.05(0.29) | -0.62, 0.52 | 0.980 |
| $LCI_{FB}$ | 3.2(1.5) % | 0.15(0.53) | -0.89, 1.19 | 0.850 |
| $Scond_{FB}$, $L^{-1}$ | - | 0.003(0.016) | -0.027, 0.033 | 0.158 |
| $Sacin_{FB}$, $L^{-1}$ | - | 0.004(0.052) | -0.098, 0.106 | 0.334 |

**FIGURE CAPTIONS**

**Figure 1. Functional residual capacity measured by controlled breathing ($FRC_{CB}$) and free breathing ($FRC_{FB}$) protocols**. (A) There was strong correlation between the protocols (r=0.94, p<0.0001). (B) Bland-Altman plot showing good agreement between the protocols (mean difference (95% limits of agreement) 0.009 (-0.666, 0.686) L, p=0.88).

**Figure 2. Functional residual capacity measured by controlled breathing ($FRC_{CB}$) and free breathing ($FRC_{FB}$) protocols versus the gold-standard body plethysmography ($FRC_{Pleth}$).** There was good correlation between FRC measured by both protocols and $FRC_{Pleth}$ (r=0.84 and r=0.92, respectively, p<0.0001 for both).

**Figure 3**. **Lung clearance index measured by controlled breathing ($LCI_{CB}$) and free breathing ($LCI_{FB}$) protocols.** (A) There was strong correlation between the protocols (r=0.84, p<0.0001). (B) Bland-Altman plot showing that free breathing produced a higher LCI compared to controlled breathing (mean difference (95% limits of agreement) 0.35 (-0.53, 1.23), p=0.0004). There was also significant proportional bias confirmed by linear regression (p=0.004).

**Figure 4. Ventilation heterogeneity in conducting airways measured by controlled breathing ($Scond_{CB}$) and free breathing ($Scond_{FB}$) protocols.** (A) There was no significant correlation between the protocols (r=0.18, p=0.36). (B) Bland-Altman plot showing high between-protocol variability (mean difference (95% limits of agreement) 0.0008 (-0.02, 0.02) $L^{-1}$, p=0.74). There appeared to be proportional bias on visual inspection, but linear regression was not statistically significant (p=0.45).

**Figure 5. Ventilation heterogeneity in distal/intra-acinar airways measured by controlled breathing ($Sacin_{CB}$) and free breathing ($Sacin_{FB}$) protocols.** (A) There was relatively poor correlation between the protocols (r=0.37, p=0.06). (B) Bland-Altman plot

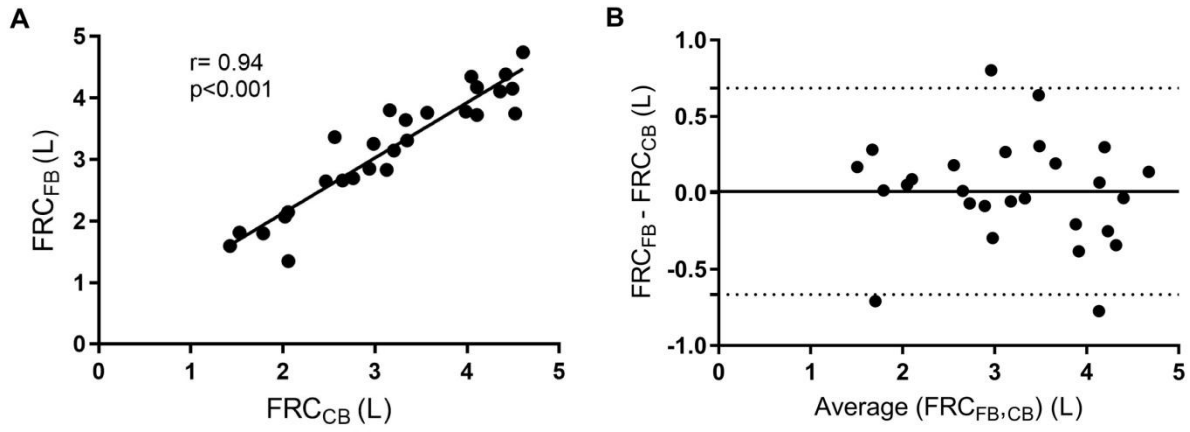
showing that free breathing produced higher Sacin compared to controlled breathing (mean difference (95% limits of agreement) 0.03 (-0.04, 0.10) $L^{-1}$, p<0.0001). There was also significant proportional bias confirmed by linear regression (p=0.002). (C) and (D) The between-protocol difference in Sacin ($Sacin_{FB}$ – $Sacin_{CB}$) was predicted by the between-protocol differences in tidal volume ($VT_{FB}$-$VT_{CB}$, regression p=0.004) and respiratory rate ($RR_{FB}$-$RR_{CB}$, regression p=0.01). n=1 participant excluded from Sacin analyses due to negative value in one trial.

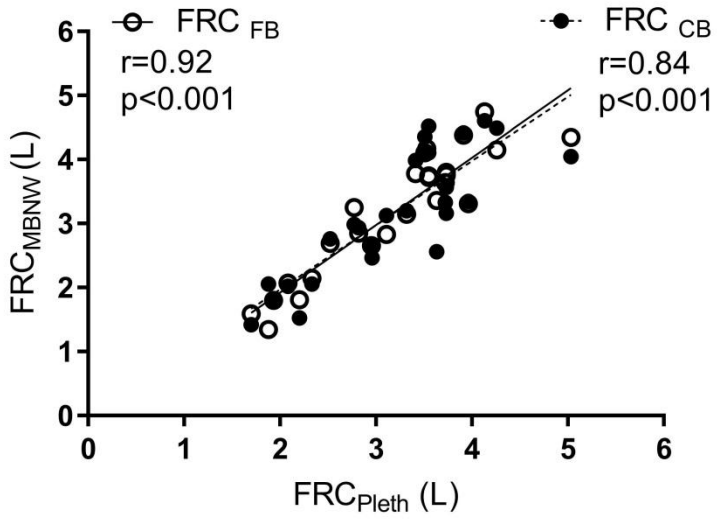
## REFERENCES

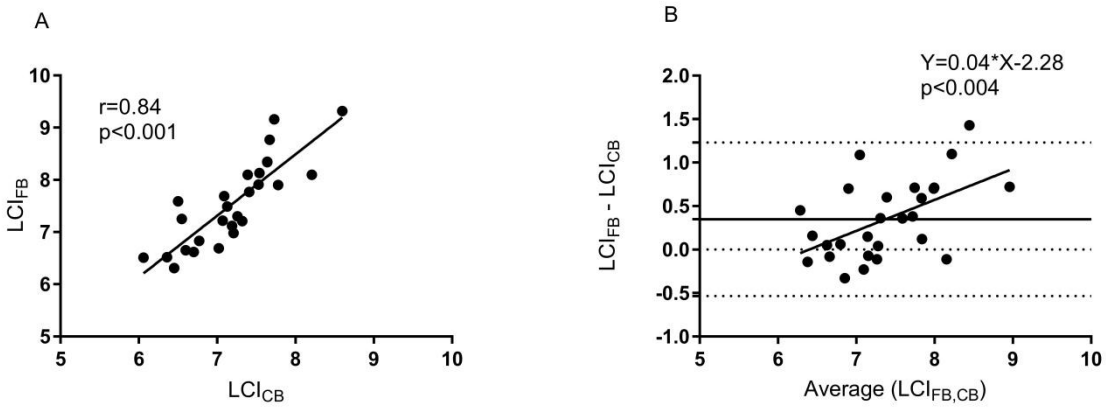

1. King GG. Cutting edge technologies in respiratory research: lung function testing. *Respirology* 2011: 16(6): 883-890.
2. Zimmermann SC, Tonga KO, Thamrin C. Dismantling airway disease with the use of new pulmonary function indices. *Eur Respir Rev* 2019: 28(151): 180122.
3. Robinson PD, Latzin P, Verbanck S, Hall GL, Horsley A, Gappa M, Thamrin C, Arets HG, Aurora P, Fuchs SI, King GG, Lum S, Macleod K, Paiva M, Pillow JJ, Ranganathan S, Ratjen F, Singer F, Sonnappa S, Stocks J, Subbarao P, Thompson BR, Gustafsson PM. Consensus statement for inert gas washout measurement using multiple- and single- breath tests. *Eur Respir J* 2013: 41(3): 507-522.
4. Crawford AB, Makowska M, Engel LA. Effect of tidal volume on ventilation maldistribution. *Respir Physiol* 1986: 66(1): 11-25.
5. Jones JG. The effect of preinspiratory lung volume on the result of the single breath O2 test. *Respir Physiol* 1967: 2(3): 375-385.
6. 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(2): 634-642.
7. Aurora P, Kozlowska W, Stocks J. Gas mixing efficiency from birth to adulthood measured by multiple-breath washout. *Respir Physiol Neurobiol* 2005: 148(1-2): 125-139.
8. Anagnostopoulou P, Latzin P, Jensen R, Stahl M, Harper A, Yammine S, Usemann J, Foong RE, Spycher B, Hall GL, Singer F, Stanojevic S, Mall MA, Ratjen F, Ramsey KA. Normative data for multiple breath washout outcomes in school-aged Caucasian children. *European Respiratory Journal* 2020: 55(4): 1901302.
9. Kjellberg S, Houltz BK, Zetterstrom O, Robinson PD, Gustafsson PM. Clinical characteristics of adult asthma associated with small airway dysfunction. *Respir Med* 2016: 117: 92-102.
10. Verbanck S, Thompson BR, Schuermans D, Kalsi H, Biddiscombe M, Stuart-Andrews C, Hanon S, Van Muylem A, Paiva M, Vincken W, Usmani O. Ventilation heterogeneity in the acinar and conductive zones of the normal ageing lung. *Thorax* 2012: 67(9): 789-795.
11. Verbanck S, Van Muylem A, Schuermans D, Bautmans I, Thompson B, Vincken W. Transfer factor, lung volumes, resistance and ventilation distribution in healthy adults. *Eur Respir J* 2016: 47(1): 166-176.

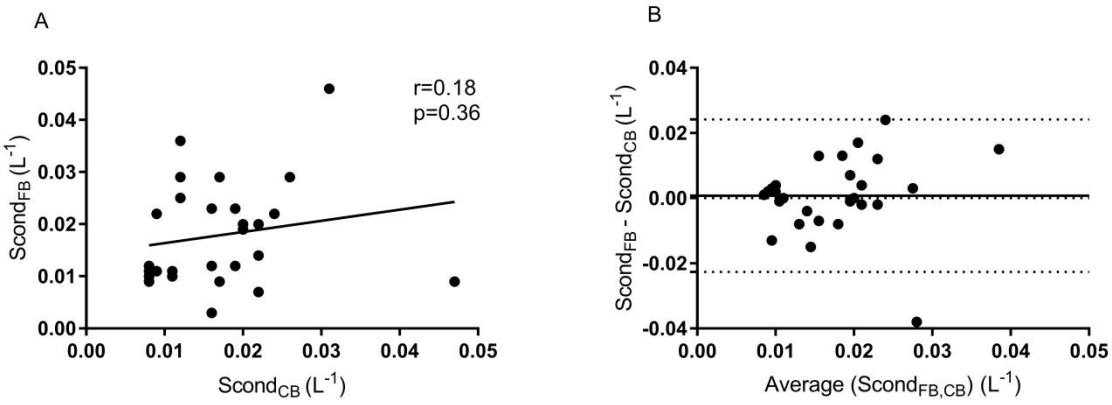
12. Yammine S, Singer F, Gustafsson P, Latzin P. Impact of different breathing protocols on multiple-breath washout outcomes in children. *J Cyst Fibros* 2014: 13(2): 190-197.
13. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MS, Zheng J, Stocks J, ERS Global Lung Function Initiative. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012: 40(6): 1324-1343.
14. Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault J-C. Lung volumes and forced ventilatory flows. *European Respiratory Journal* 1993: 6(Suppl 16): 5-40.
15. Verbanck S, Schuermans D, Van Muylem A, Paiva M, Noppen M, Vincken W. Ventilation distribution during histamine provocation. *J Appl Physiol* 1997: 83(6): 1907-1916.
16. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* 2016: 15(2): 155-163.
17. Tonga KO, Robinson PD, Farah CS, King GG, Thamrin C. In vitro and in vivo functional residual capacity comparisons between multiple-breath nitrogen washout devices. *ERJ Open Research* 2017: 3(4).
18. Poncin W, Singer F, Aubriot A-S, Lebecque P. Agreement between multiple-breath nitrogen washout systems in children and adults. *Journal of Cystic Fibrosis* 2017: 16(2): 258-266.
19. Christensson P, Arborelius M, Jr., Kautto R. Volume of trapped gas in lungs of healthy humans. *J Appl Physiol Respir Environ Exerc Physiol* 1981: 51(1): 172-175.
20. Edward JBJ, Macklem. P. Airway Closure: Demonstration by breathing 100% $O_2$ at low lung volumes and by $N_2$ washout. *J Appl Physiol* 1968: 25(2): 139-148.
21. Stuart-Andrews CR, Kelly VJ, Sands SA, Lewis AJ, Ellis MJ, Thompson BR. Automated detection of the phase III slope during inert gas washout testing. *J Appl Physiol* 2012: 112(6): 1073-1081.
22. Bates JHT, Peters U. A model-based approach to interpreting multibreath nitrogen washout data. *J Appl Physiol* 2018: 124(5): 1155-1163.
23. Lewis SM, Evans JW, Jalowayski AA. Continuous distributions of specific ventilation recovered from inert gas washout. *Journal of Applied Physiology* 1978: 44(3): 416-423.
24. Ratjen F, Jensen R, Klingel M, McDonald R, Moore C, Benseler N, Wilson D, Stanojevic S. Effect of changes in tidal volume on multiple breath washout outcomes. *PLoS One* 2019: 14(7): e0219309.
25. Singer F, Houltz B, Latzin P, Robinson P, Gustafsson P. A realistic validation study of a new nitrogen multiple-breath washout system. *PloS one* 2012: 7(4): e36083-e36083.
26. Zwitserloot AM, van den Born EJ, Raaijmakers LHA, Stalman WE, van Smaalen M, van den Berge M, Gappa M, Koppelman GH, Willemse BWM. Differences in lung clearance index and functional residual capacity between two commercial multiple-breath nitrogen washout devices in healthy children and adults. *ERJ Open Research* 2020: 6(2): 00247-02019.
27. Bell AS, Lawrence PJ, Singh D, Horsley A. Feasibility and challenges of using multiple breath washout in COPD. *International journal of chronic obstructive pulmonary disease* 2018: 13: 2113-2119.
28. Horsley AR, Macleod KA, Robson AG, Lenney J, Bell NJ, Cunningham S, Greening AP, Gustafsson PM, Innes JA. Effects of cystic fibrosis lung disease on gas mixing indices derived from alveolar slope analysis. *Respiratory Physiology & Neurobiology* 2008: 162(3): 197-203.
29. Isaac SM, Jensen R, Anagnostopoulou P, Davies JC, Gappa M, Latzin P, Saunders C, Short C, Singer F, Stanojevic S, Zwitserloot A, Ratjen F. Evaluation of a multiple breath nitrogen washout system in children. *Pediatric Pulmonology* 2020: 55(8): 2108-2114.

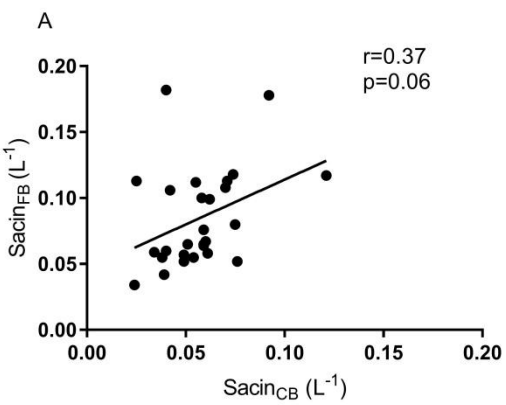
30. Gonem S, Corkill S, Singapuri A, Gustafsson P, Costanza R, Brightling CE, Siddiqui S. Between-visit variability of small airway obstruction markers in patients with asthma. *Eur Respir J* 2014: 44(1): 242-244.
31. Jetmalani K, Chapman DG, Thamrin C, Farah CS, Berend N, Salome CM, King GG. Bronchodilator responsiveness of peripheral airways in smokers with normal spirometry. *Respirology* 2016: 21(7): 1270-1276.
32. Verbanck S, Paiva M. Model simulations of gas mixing and ventilation distribution in the human lung. *J Appl Physiol* 1990: 69(6): 2269-2279.
33. Fowler WS. Lung function studies. III. Uneven pulmonary ventilation in normal subjects and in patients with pulmonary disease. *J Appl Physiol* 1949: 2(6): 283-299.

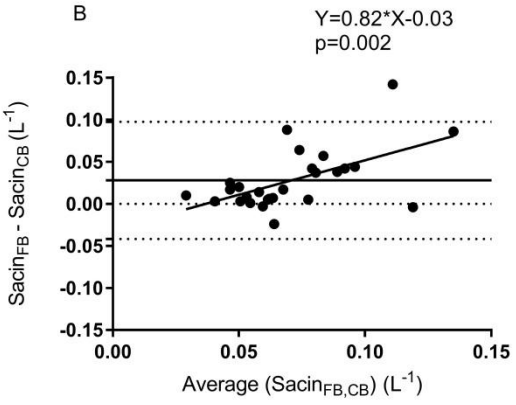

A
r= 0.94
p<0.001
FRC_FB (L)
FRC_CB (L)
0
1
2
3
4
5
B
FRC_FB - FRC_CB (L)
Average (FRC_FB,CB) (L)
1.0
0.5
0.0
-0.5
-1.0

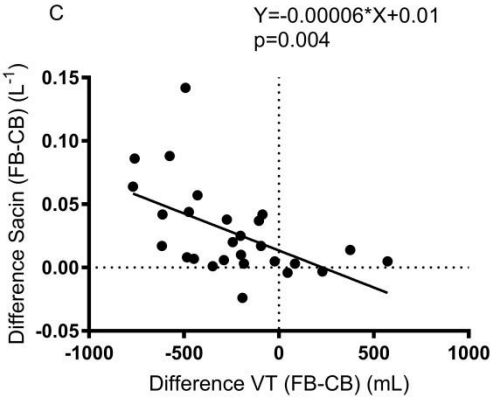

FRC FB
r=0.92
p<0.001
FRC CB
r=0.84
p<0.001
FRCMBNW (L)
FRCPleth (L)
0
1
2
3
4
5
6

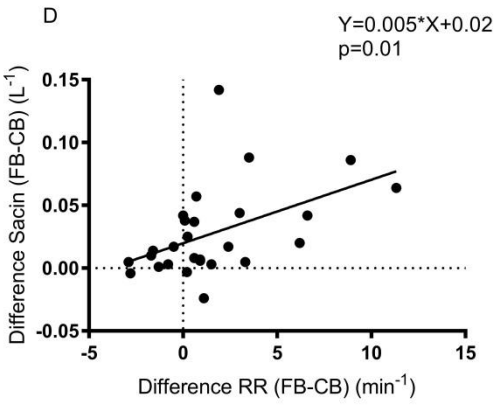

A
r=0.84
p<0.001
LCI_FB
LCI_CB
B
Y=0.04*X-2.28
p<0.004
LCI_FB - LCI_CB
Average (LCI_FB,CB)


A
r=0.18
p=0.36
Scond$_{FB}$ (L$^{-1}$)
Scond$_{CB}$ (L$^{-1}$)
0.00
0.01
0.02
0.03
0.04
0.05
B
Scond$_{FB}$ - Scond$_{CB}$ (L$^{-1}$)
Average (Scond$_{FB,CB}$) (L$^{-1}$)
-0.04
-0.02
0.00
0.02
0.04


A
r=0.37
p=0.06
Sacin$_{FB}$ (L$^{-1}$)
Sacin$_{CB}$ (L$^{-1}$)
0.00
0.05
0.10
0.15
0.20


B
Y=0.82*X-0.03
p=0.002
Sacin$_{FB}$ - Sacin$_{CB}$ (L$^{-1}$)
Average (Sacin$_{FB,CB}$) (L$^{-1}$)
-0.15
-0.10
-0.05
0.00
0.05
0.10
0.15


C
Y=-0.00006*X+0.01
p=0.004
Difference Sacin (FB-CB) (L$^{-1}$)
Difference VT (FB-CB) (mL)
-1000
-500
0
500
1000
-0.05
0.00
0.05
0.10
0.15


D
Y=0.005*X+0.02
p=0.01
Difference Sacin (FB-CB) (L$^{-1}$)
Difference RR (FB-CB) (min$^{-1}$)
-5
0
5
10
15
-0.05
0.00
0.05
0.10
0.15

# Controlled vs free breathing for multiple breath nitrogen washout in healthy adults

## Online Supplement

**Blake M HANDLEY[1,2], Edward JEAGAL[2], Robin E SCHOEFFEL[1], Tanya BADAL[2], David G CHAPMAN[2,3], Catherine E FARROW[2,6], Gregory G KING[1,2,4], Paul D ROBINSON[5], Stephen MILNE[4,7]*, Cindy THAMRIN[2,4]***

1. Department of Respiratory Medicine, Royal North Shore Hospital, St Leonards, NSW, Australia
2. Woolcock Institute of Medical Research, University of Sydney, Glebe NSW, Australia
3. Translational Airways Group, School of Life Sciences, Faculty of Science, University of Technology Sydney, Ultimo NSW, Australia
4. Faculty of Medicine and Health, University of Sydney, Camperdown NSW, Australia
5. Department of Respiratory Medicine, The Children's Hospital at Westmead, Westmead NSW, Australia.
6. Department of Respiratory and Sleep Medicine, Westmead Hospital, Westmead NSW, Australia
7. Centre for Heart Lung Innovation, St Paul's Hospital and Division of Respiratory Medicine, University of British Columbia, Vancouver BC, Canada

***joint Senior Authors with equal contributions**

**CORRESPONDING AUTHOR:**

Dr Cindy Thamrin
Airway Physiology and Imaging Group
Woolcock Institute of Medical Research
431 Glebe Point Rd
Glebe NSW 2037
AUSTRALIA
E: cindy.thamrin@sydney.edu.au

**CONTENTS**

## 1. METHODS: Participant inclusion and exclusion criteria

Inclusion:

- >18 years of age
- Free of respiratory disease
- Free of cardiovascular conditions
- No current respiratory symptoms
- No regular use of respiratory medications

Exclusion:

- Current smoking
- ≥5 pack years past smoking history
- Symptoms of respiratory tract infection in the previous 6 weeks
- Unable to provide informed written consent

## 2. METHODS: MBNW testing protocols

We used the Exhalyzer D with Spiroware v3.1.6 (Eco Medics AG, Duernten, Switzerland). This device measures flow via a mainstream ultrasonic flowmeter, oxygen ($O_2$) concentration by a side-stream laser sensor, and carbon dioxide ($CO_2$) by a main-stream infra-red sensor. The device measures nitrogen ($N_2$) concentration indirectly by subtraction of $CO_2$, argon and $O_2$ concentrations. The software accounted for the pre- and post-gas-sampling dead space (47 mL and 22 mL, respectively), and BTPS corrections.

*Device calibration and quality control*

Prior to each testing session, the flow sensor was calibrated using a 1 L syringe, and gas analysers were calibrated using medical air and 100 percent O2. Periodically (weekly) during the study period, gas and flow signal synchronisation was performed by a control operator breathing on the mouth piece. As a functional calibration, an “octopus” syringe lung model of known “functional residual capacity (FRC)” [1] was also tested periodically (1-4 weeks) to ensure the measured FRC was within

an acceptable range. RNSH and WIMR laboratories both conduct internal biological control programmes where 2-3 nominated healthy individuals undergo testing on a monthly basis, or as required following changes in conditions (moving of equipment, replacement of parts, software updates, etc). A standard control chart approach was used, whereby deviations from the baseline average of greater than 2xSDs were considered significant cause for investigation. There were no deviations from the expected ranges during the testing period.

*Test procedure*

Volunteers were asked to sit in an upright, comfortable position with bite-on rubber mouthpiece positioned securely within the mouth, whilst maintaining a tight seal with the lips and a neutral head position. The operator monitored the test using real-time flow/volume and volume/time traces. All trials started with a period of normal relaxed breathing on room air in order to establish a stable EELV as indicated on the volume-time trace. After determining that a stable breathing pattern on room air had been established, the operator switched the breathing circuit to 100% $O_2$. The operator visually monitored the $N_2$ concentration and volume/time traces to ensure there were no mouth leaks, coughs, or inspirations that exceeded the bias flow. If these were present, the operator terminated the trial and repeated it after the appropriate wash-in time (twice the length of the previous washout [2]). When the mean $N_2$ concentration was $1/40^{th}$ of the initial concentration, the operator asked the participant to breathe a further 5-6 breaths before terminating the trial. We considered a measurement session to be complete when there were at least 3 technically acceptable trials (i.e technically acceptable $1^{st}$ breath, at least 2/3 of total breaths technically acceptable, end of trial criteria met, and absence of artefacts), as per the current consensus statement [2] with FRC values within ±10% of the mean of the 3 trials.

For the free breathing protocol, the operator instructed the participant to “breathe relaxed and normally” on room air. The software’s visual incentive screen was switched off at all times during this protocol. After an initial period of breathing stabilisation (approximately 30 s), the operator commenced the washout by switching the circuit to 100% $O_2$ and the participant continued to breathe normally for the duration of the trial. Sometimes, at the time of acquisition, the operator observed the participant was not breathing sufficiently deeply for an adequate phase III slope to be captured. In

those cases, the operator requested that the participant breathed “a little deeper” and noted this instruction in the participant’s file.

For the controlled breathing protocol, after the initial breathing stabilisation period, the operator instructed the participant to breathe at a VT of 0.95 – 1.3 L and at RR 8-12/min with the use of the visual incentive screen as originally described by (described by Verbanck et al [3]);. Once this was achieved, and the operator was satisfied that a stable end-expiratory lung volume had been reached, they commenced the washout by switching the circuit to 100% $O_2$. The participant continued this same pattern of breathing until the end of the trial.

## 3. METHODS: MBNW analysis

We analysed MBNW data using Spiroware software (v.3.1.6).

*$N_2$ Phase III slope (SIII) calculation*

Breaths for which the operator decided there was insufficient Phase III to accurately estimate the slope were excluded. By default, the software estimates SIII as the linear regression between 50-95% of expired volume. However, the operator manually adjusted the boundaries for SIII determination where needed, e.g. to exclude Phase II (particularly when expired volume was relatively small) and when there were prominent cardiogenic oscillations.

While the test operator performed preliminary analysis on individual trials during the testing session, a single investigator re-analysed the data post hoc for all participants as a batch in order to ensure a consistent approach to analysis.

*Calculation of MBNW indices*

MBNW indices are calculated automatically by the Spiroware software.

- FRC is calculated as the ratio of exhaled $N_2$ volume to the difference in initial and final end-tidal concentrations.
- Phase III slopes, normalised for mean expired $N_2$ concentration within phase III for that breath ($S_{nIII}$), are plotted as a function of lung turnover (i.e. cumulative expired volume [CEV] divided by FRC).

- LCI is calculated by dividing the CEV measured at 1/40$^{th}$ of initial $N_2$ concentration by FRC [2].
- Scond is calculated as the slope of a linear regression of $S_{nIII}$ between the limits of 1.5 and 6 lung turnovers.
- Sacin is calculated as $S_{nIII}$ of the first breath minus Scond.

For the free breathing protocol, to allow comparison of Scond and Sacin between participants with different lung sizes and breathing at different VT, each SnIII is divided by FRC and then multiplied by FRC*VT of the breath, resulting in a net multiplication by VT [4]. These adjusted indices are denoted in the literature and Spiroware software as Scond*VT and Sacin*VT; however, in keeping with the nomenclature described above, we refer to them simply as $Scond_{FB}$ and $Sacin_{FB}$, respectively.

## 4. RESULTS: Table S1 – MBNW ventilation heterogeneity from free breathing (Z scores)

| | Whole group | Repeatability subgroup |
|---|---|---|
| *Free breathing* | | |
| $Z\text{-}LCI_{FB}$ | 1.56(1.73) | 1.38(1.79) |
| $Z\text{-}Scond_{FB}$ | -0.2(2.22) | 0.12(1.76) |
| $Z\text{-}Sacin_{FB}$ | 1.15(1.88) | 0.03(1.01) |

Values are mean(SD). For the repeatability subgroup, results are from the first visit. Reference equations for Z-scores from the free breathing protocol were from Kjellberg 2016 [4]. These equations were derived from data collected on the same MBNW device used in the present study. LCI, lung clearance index; Scond, conductive zone ventilation heterogeneity; Sacin, acinar zone ventilation heterogeneity.

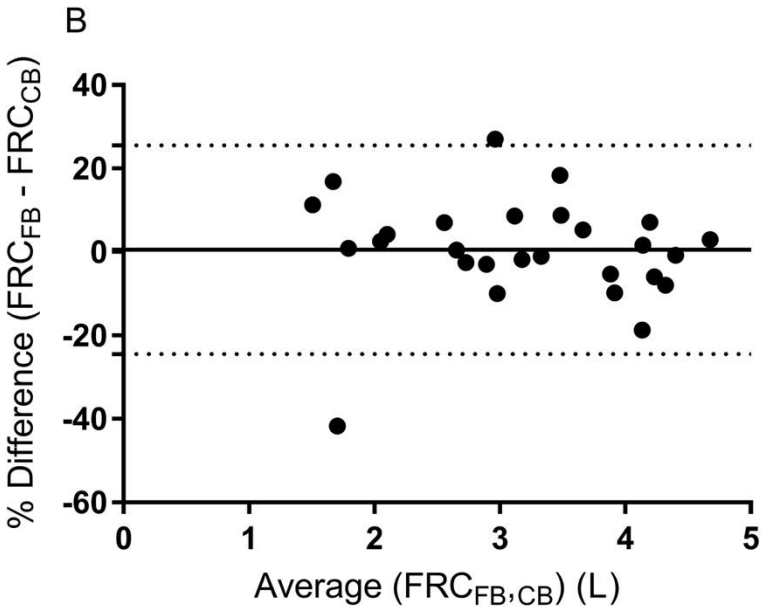
## 5. RESULTS: Table S2 – Scond and Sacin before and after VT correction

| | Controlled Breathing | Free Breathing |
|---|---|---|
| Scond | 0.017(0.009) | 0.022(0.022) |
| Scond*VT | 0.019(0.009) | 0.018(0.01) |
| Sacin | 0.057(0.022) | 0.13(0.125) |
| Sacin*VT | 0.064(0.023) | 0.085(0.038) |

Values are mean(SD). During the free breathing protocol, a correction for VT is typically made to allow comparison of Scond and Sacin between individuals with different lung sizes and breathing at different VT. These corrected indices are denoted Scond*VT and Sacin*VT. In line with the current consensus recommendations (Robinson et al 2013) [2], we are reporting both corrected and uncorrected values for both protocols. (Note: in the rest of the manuscript, Scond*VT and Scain*VT from the free breathing protocol are referred to as $Scond_{FB}$ and $Sacin_{FB}$, wheras $Scond_{CB}$ and $Sacind_{CB}$ refer to the uncorrected values from the controlled breathing protocol).

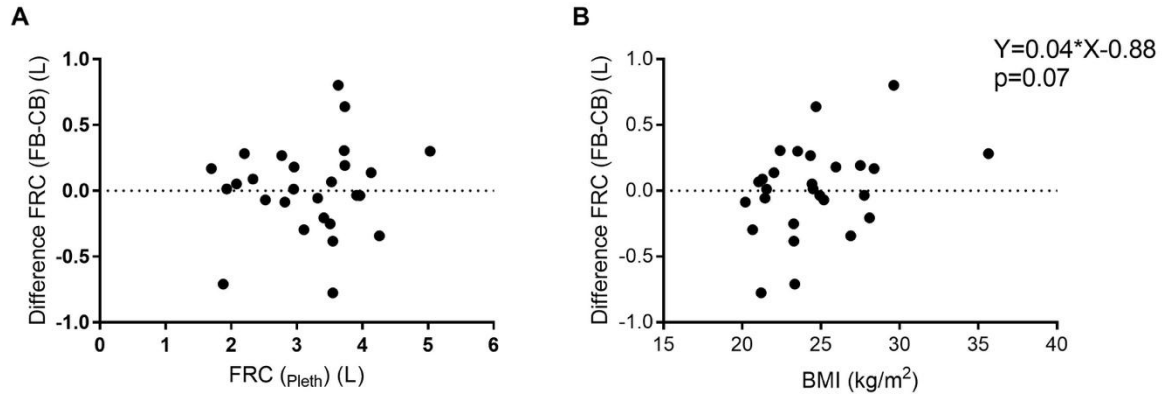
## 6. RESULTS: Figure S1

**Figure S1. Percent difference in functional residual capacity measured by controlled breathing and free breathing protocols.** Bland-Altman plot showing good agreement between the protocols (mean difference (95% limits of agreement) 0.54 (-24.43, 25.51) %, p=0.76).



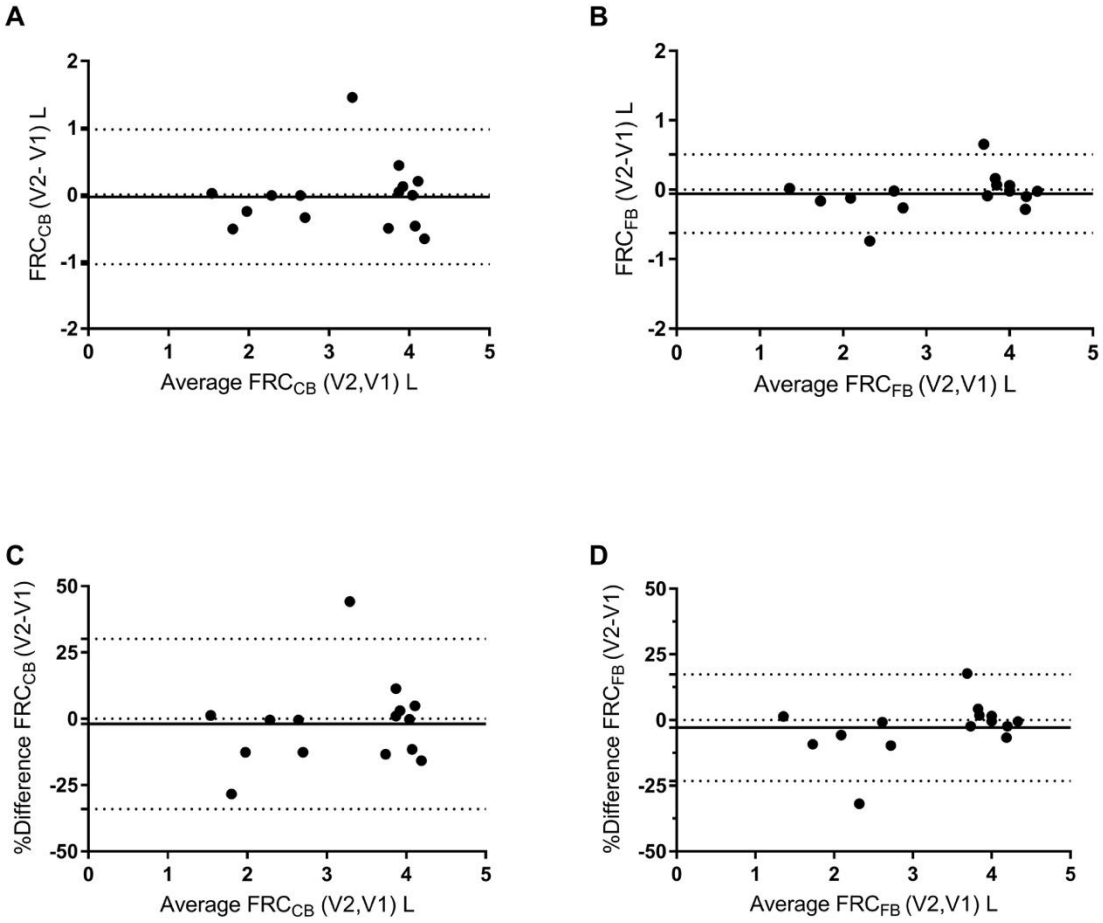
## 7. RESULTS: Figure S2

**Figure S2. Predictors of differences in functional residual capacity measured by controlled breathing and free breathing MBNW protocols.** Between-protocol differences were not predicted by FRC measured by the gold-standard body plethysmography (linear regression p=0.55) (A), but may be related to the individual's body mass index (linear regression of borderline significance, p=0.07) (B).



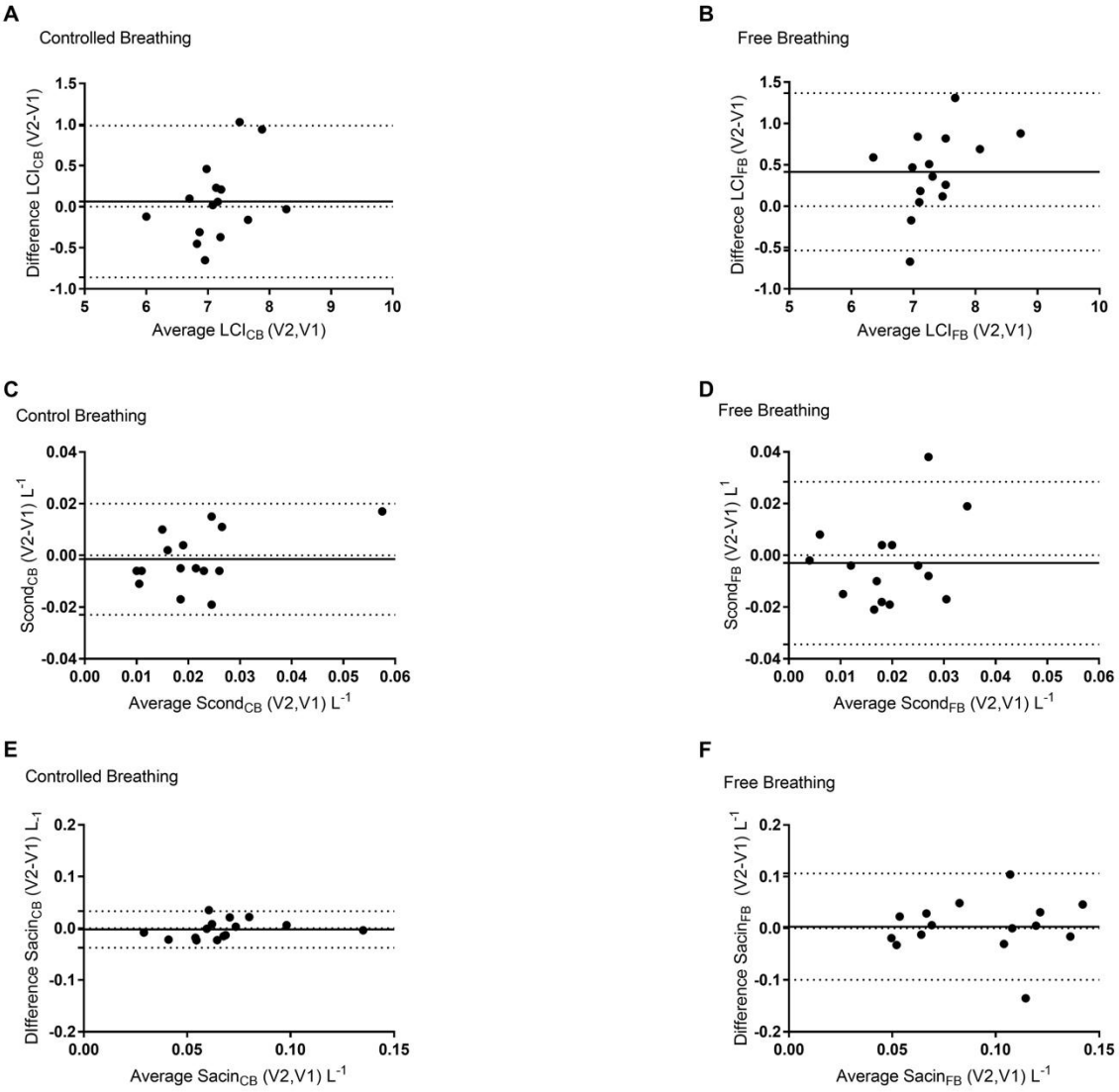
## 8. RESULTS: Figure S3

**Figure S3. Bland-Altman plots of between-session difference (Visit 2 minus Visit 1) for functional residual capacity measured by controlled breathing (FRCCB, left panels) and free breathing (FRCFB, right panels) MBNW protocols.** Mean absolute difference (95% limits of agreement [LOA]) for (A) controlled breathing -0.03 (-1.04, 0.98) L and (B) free breathing -0.05 (-0.62, 0.52). Mean percent difference (95% LOA) for (C) controlled breathing -1.98 (-34.05, 30.1) % and (D) free breathing -2.89 (-23.18, 17.41) %.



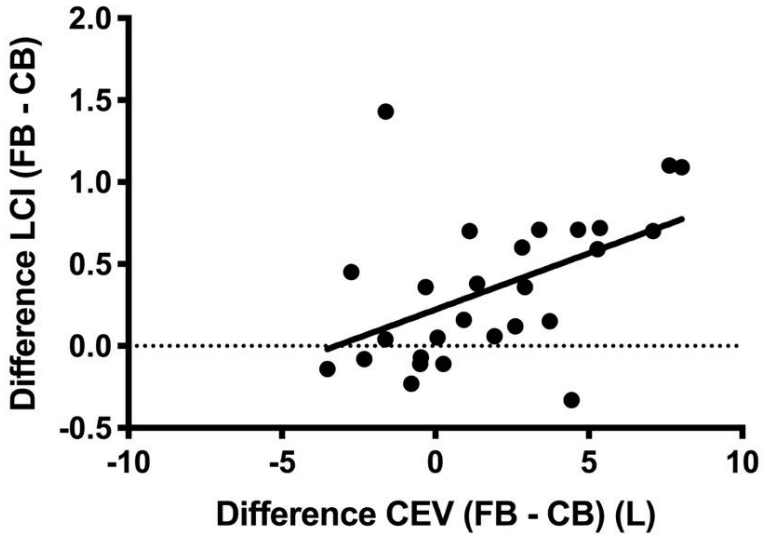
## 9. RESULTS: Figure S4

**Figure S4: Bland-Altman plots of between-session difference (Visit 2 minus Visit 1) for ventilation heterogeneity indices measured by controlled breathing (left panels) and free breathing (right panels) MBNW protocols.** Values are mean difference (95% limits of agreement). Lung clearance index (A) controlled breathing 0.06 (-0.86, 0.98) and (B) free breathing 0.42 (-0.54, 1.37). Scond (C) controlled breathing -0.001 (-0.017, 0.015) and (D) free breathing 0.003 (-0.027, 0.033). Sacin (E) controlled breathing -0.003 (-0.037, 0.031) and (F) free breathing 0.004 (-0.098, 0.106).



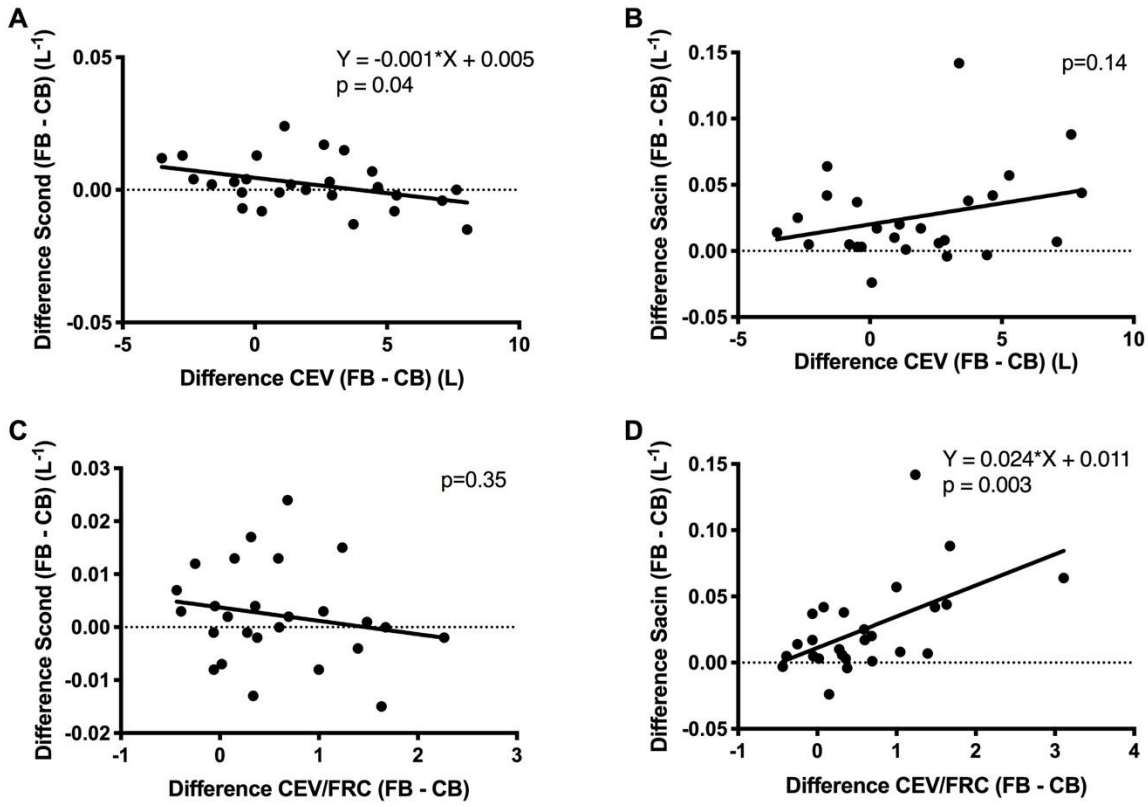
## 10. RESULTS: Figure S5

**Figure S5: Relationship between differences in functional residual capacity (FRC) and change in cumulative expired volume (CEV) between controlled breathing and free breathing protocols.**



## 11. RESULTS: Figure S6

**Figure S5: Between-protocol differences in Scond and Sacin plotted against differences in cumulative expired volume (CEV, top panels) and CEV corrected for functional residual capacity (FRC, bottom panels).** Between-protocol difference in Scond was related to between-protocol difference in CEV (linear regression p=0.04) (A) but not CEV/FRC (p=0.35) (C). Between-protocol difference in Sacin was not related to difference in CEV (p=0.14) (B) but was significantly related to CEV/FRC (p=0.003) (D). n=1 outlier excluded from Scond analyses. n=1 participant excluded from Sacin analyses due to negative value in one trial.



## 12. REFERENCES


1. Tonga KO, Robinson PD, Farah CS, King GG, Thamrin C. In vitro and in vivo functional residual capacity comparisons between multiple-breath nitrogen washout devices. *ERJ Open Research* 2017: 3(4).
2. Robinson PD, Latzin P, Verbanck S, Hall GL, Horsley A, Gappa M, Thamrin C, Arets HG, Aurora P, Fuchs SI, King GG, Lum S, Macleod K, Paiva M, Pillow JJ, Ranganathan S, Ratjen F, Singer F, Sonnappa S, Stocks J, Subbarao P, Thompson BR, Gustafsson PM. Consensus statement for inert gas washout measurement using multiple- and single- breath tests. *Eur Respir J* 2013: 41(3): 507-522.
3. Verbanck S, Schuermans D, Van Muylem A, Paiva M, Noppen M, Vincken W. Ventilation distribution during histamine provocation. *J Appl Physiol* 1997: 83(6): 1907-1916.
4. Kjellberg S, Houltz BK, Zetterstrom O, Robinson PD, Gustafsson PM. Clinical characteristics of adult asthma associated with small airway dysfunction. *Respir Med* 2016: 117: 92-102.