The effect of oxygen and carbon dioxide cross-sensitivity sensor error in the Eco Medics Exhalyzer D device on measures of conductive and acinar airway function

Jack Bozier1,2, Edward Jeagal1,3, Paul D. Robinson1,4, G. Kim Prisk1,5, David G. Chapman 1,2,3, Gregory G. King1,2,6, Cindy Thamrin 1,6 and Sandra Rutting 1,2

1Woolcock Institute of Medical Research, The University of Sydney, Glebe, NSW, Australia. 2Dept of Respiratory Medicine, Royal North Shore Hospital, St Leonards, NSW, Australia. 3School of Life Sciences, University of Technology Sydney, Glebe, NSW, Australia. 4Dept of Respiratory Medicine, The Children’s Hospital at Westmead, Westmead, NSW, Australia. 5University of California, San Diego, CA, USA. 6Faculty of Medicine and Health, The University of Sydney, Glebe, NSW, Australia.

Corresponding author: Sandra Rutting (sandra.rutting@sydney.edu.au)

Shareable abstract (@ERSpublications)
O2 and CO2 cross-sensitivity sensor error in the Exhalyzer D device significantly overestimates FRC and LCI in adults, consistent with infants and children. Importantly, there was a high degree of underestimation of $S_{cond}$ but minimal impact on $S_{acin}$.

https://bit.ly/3HcH3Tp

Cite this article as: Bozier J, Jeagal E, Robinson PD, et al. The effect of oxygen and carbon dioxide cross-sensitivity sensor error in the Eco Medics Exhalyzer D device on measures of conductive and acinar airway function. ERJ Open Res 2022; 8: 00614-2021 [DOI: 10.1183/23120541.00614-2021].

Abstract

Introduction The multiple breath nitrogen washout (MBNW) test provides important clinical information in obstructive airways diseases. Recently, a significant cross-sensitivity error in the O2 and CO2 sensors of a widely used commercial MBNW device (Exhalyzer D, Eco Medics AG, Duernten, Switzerland) was detected, which leads to overestimation of N2 concentrations. Significant errors in functional residual capacity (FRC) and lung clearance index (LCI) have been reported in infants and children. This study investigated the impact in adults, and on additional important indices reflecting conductive ($S_{cond}$) and acinar ($S_{acin}$) ventilation heterogeneity, in health and disease.

Methods Existing MBNW measurements of 27 healthy volunteers, 20 participants with asthma and 16 smokers were reanalysed using SPIROWARE V 3.3.1, which incorporates an error correction algorithm. Uncorrected and corrected indices were compared using paired t-tests and Bland–Altman plots.

Results Correction of the sensor error significantly lowered FRC (mean difference 9%) and LCI (8–10%) across all three groups. $S_{cond}$ was higher following correction (11%, 14% and 36% in health, asthma and smokers, respectively) with significant proportional bias. $S_{acin}$ was significantly lower following correction in the asthma and smoker groups, but the effect was small (2–5%) and with no proportional bias.

Discussion The O2 and CO2 cross-sensitivity sensor error significantly overestimated FRC and LCI in adults, consistent with data in infants and children. There was a high degree of underestimation of $S_{cond}$ but minimal impact on $S_{acin}$. The presence of significant proportional bias indicates that previous studies will require reanalysis to confirm previous findings and to allow comparability with future studies.
MBNW has been extensively used as a research tool in various respiratory diseases, particularly in obstructive airways diseases. With the availability of commercially available devices and international guidelines, it has emerging utility in clinical care, especially in cystic fibrosis (CF). The LCI has proved to be a sensitive marker of early disease progression in children with CF and has also been included as a primary end-point in several therapeutic trials [4, 5]. MBNW has yet to be a part of clinical management in other lung diseases, but studies have shown utility of $S_{\text{acin}}$ and $S_{\text{cond}}$ in guiding up- versus down-titration of treatment [6, 7] and sensitivity to detect improvement in symptoms in response to treatment with high-dose inhaled corticosteroid [8] or monoclonal antibody therapy in asthma [9]. These indices are also sensitive markers of small airway dysfunction and its reversibility in smokers with normal spirometry [10, 11].

Recently, the presence and impact of a critical sensor error in a commercial device used to perform MBNW (Exhalyzer D, Eco Medics AG, Duernten, Switzerland) has been reported in infants and older children [12, 13]. This MBNW device relies on accurate measurements from $\text{O}_2$ and $\text{CO}_2$ sensors to calculate $\text{N}_2$ concentration indirectly. It was found that both sensors exhibit cross-sensitivities, i.e. the $\text{O}_2$ sensor estimation is dependent on $\text{CO}_2$ concentrations and vice versa, such that as the washout progresses, $\text{O}_2$ and $\text{CO}_2$ concentrations are underestimated and $\text{N}_2$ concentrations increasingly overestimated, prolonging the washout. This has been shown to result in significant errors of up to 12% and 15–19% in the assessment of FRC and LCI, respectively [12–14]. A software update (V 3.3.1) has now been released by the manufacturer with an implemented correction algorithm, which recalculates the $\text{N}_2$ concentration trace.

The magnitude of effect of this sensor error correction on these MBNW indices in adults is currently unknown, and to date there has been no description of the effects on additional important indices such as $S_{\text{cond}}$ and $S_{\text{acin}}$. This is essential to understand the validity of changes reported in previously published studies. Therefore, this study aimed to determine the effect of the $\text{CO}_2$ and $\text{O}_2$ sensor correction on MBNW parameters in both health and disease by examining three different adult cohorts: 1) healthy volunteers, 2) patients with asthma, and 3) long-term smokers. Secondly, we investigated whether correction of the sensor error affected the within- and between-session repeatability of MBNW parameters in health. Some of the data from the healthy and asthma participants have been previously published [15, 16].

**Methods**

**Research participants**

In this study we retrospectively reanalysed MBNW measurements from healthy volunteers, participants with asthma and long-term smokers that were recruited from Royal North Shore Hospital and the Woolcock Institute of Medical Research. Healthy participants were current nonsmokers with a smoking history of <10 pack-years and no respiratory disease. Patients with asthma had a physician diagnosis of asthma and were current nonsmokers with a smoking history of <10 pack-years. Long-term smokers were current smokers with at least a 10 pack-year smoking exposure; these data were collected as part of a larger clinical trial (Australian Clinical Trials Registration Number (ACTRN): 1261601208493) in smokers with normal post-bronchodilator (BD) spirometry or GOLD Stage 1 (post-BD FEV1/FVC <0.7 but FEV1 >80% predicted), with the additional inclusion criteria of abnormal $S_{\text{cond}}$ and/or $S_{\text{acin}}$ as assessed by z-score $<-1.64$ using published predicted equations [11]. The original studies were approved by the local Human Research Ethics Committee (Northern Sydney Local Health District, LNR/16/HAWKE/11 and HREC/15/HAWKE/489).

**Standard pulmonary function testing**

After obtaining written informed consent, all participants underwent conventional lung function testing including spirometry, plethysmography and diffusing capacity for carbon monoxide ($D_{\text{CO}}$). These were performed according to American Thoracic Society (ATS)/European Respiratory Society (ERS) criteria. All parameters were expressed as percent predicted using published predicted equations [17, 18].

**MBNW testing**

In the original studies, after a period of at least 10 min of rest, the healthy and asthmatic participants underwent MBNW testing by two commonly used breathing protocols: controlled and free-breathing protocols, in randomised order (assigned by a computer-based random number generator); the group of smokers performed MBNW using the controlled breathing protocol only. A subset of healthy participants returned for testing within 3 months of their first visit, in which all measurements were repeated in the same order. Both controlled and free-breathing protocols were included as several published studies showed that indices of conductive and acinar ventilation heterogeneity were not comparable between breathing protocols [15, 16, 19].
MBNW was performed using the Exhalyzer D with SPIROWARE V 3.1.6 (Eco Medics AG, Duemiten, Switzerland). Both the controlled breathing and free-breathing protocols were performed according to ERS consensus and have been previously described in detail [15, 20]. In brief, after establishing a stable breathing pattern and end-expiratory lung volume (EELV), nitrogen washout during 100% O2 inhalation was commenced. The controlled breathing protocol required participants to breathe at a RR between 8 and 12 breaths.min$^{-1}$ and tidal volume ($V_T$) between 0.95 and 1.3 L following visual feedback until the N2 concentration decreased to 1/40th of the starting end-expiratory N2 concentration. In the free-breathing protocol, participants were encouraged to adopt relaxed tidal breathing but advised to adjust tidal volumes if insufficient expired N2 phase III slope was observed; calculated $S_{\text{cond}}$ and $S_{\text{acin}}$ were adjusted for $V_T$, as per consensus guidelines [20]. At least three technically acceptable trials with FRC values <10% of the mean were obtained for each breathing protocol.

**MBNW analysis**

The effect of the sensor error was assessed by comparing the parameters of standard (uncorrected) analysis in SPIROWARE V 3.1.6 with corrected parameters reanalysed in new SPIROWARE V 3.3.1, applying the sensor error correction algorithm. The correction algorithm has been described extensively before in Sandvik et al. [13] and Wyler et al. [12]. Briefly, the algorithm was derived using Exhalyzer D sensors and mass spectrometer to measure the O2 and CO2 concentrations of a wide range of well-defined technical gas mixtures under various conditions, and used a polynomial function to correct for the errors observed. System settings, delay correction and quality control remained unaltered (i.e. selection of breaths and any correction made to phase III slopes were consistent between both versions).

**Statistical analysis**

Statistical analysis was carried out with GraphPad Prism 8 (GraphPad Software Inc., La Jolla, CA, USA). All data are expressed as mean±SD, unless otherwise stated. Differences between uncorrected and corrected parameters were examined using paired Student’s t-tests and Pearson’s correlation. To investigate bias, we generated Bland–Altman plots as the difference (corrected minus uncorrected) versus the average, plotting the mean difference and 95% limit of agreement (95% LoA). We performed linear regression of the difference versus average to determine any proportional bias.

To make clear the consequence of the correction of the sensor error on prior studies, we present these results as the change in the outcome parameters of existing studies that result from this correction, i.e. with the uncorrected parameters as reference (for example, the sensor error results in expired N2 being erroneously high towards the end of the washout). This in turn causes an overestimation in FRC. Our results are presented in the context of how FRC is altered when the sensor error is corrected, in this case a reduction in calculated FRC.

Within-session variability was expressed as the coefficient of variation (CoV) calculated as the ratio of the SD to the mean from three separate trials. To determine between-session variability, we calculated the difference (visit 2 minus visit 1) and 95% LoA separately for corrected and uncorrected parameters. 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). A p-value below 0.05 was considered statistically significant.

**Results**

**Patient demographics**

We reanalysed MBNW measurements from 27 healthy volunteers, 20 asthmatic patients and 16 long-term smokers. The patients’ demographics and lung function are summarised in table 1. The healthy volunteers were slightly younger than the asthmatic patients and smokers. The group of smokers had a mean±SD smoking history of 19.3±8.6 pack-years. Both plethysmography and MBNW-derived FRC were comparable across the groups, whereas MBNW indices of heterogeneity were significantly higher in the asthma and smoker groups compared to health, and higher in the smokers compared to asthma (in terms of $S_{\text{cond}}$ and $S_{\text{acin}}$).

**Effects of sensor correction on MBNW parameters**

Correction of CO2 and O2 sensor error had a significant effect on all MBNW parameters measured by the controlled breathing protocol (table 2). Following correction, mean (95% CI) FRC and LCI decreased by 7.8 (7.0–8.4)% and 9.8 (8.8–10.8)% respectively, in health. Similar decreases in FRC and LCI were observed in asthma and long-term smokers. While uncorrected FRC values measured by MBNW were comparable to FRC measured by body plethysmography, corrected FRC values were significantly lower.
TABLE 1  As participant demographics and lung function

| Parameter          | Health (n=27) | Asthma (n=20) | Smokers (n=16) |
|--------------------|---------------|---------------|---------------|
| Participants (n)   | 27            | 20            | 16            |
| Males/females (n)  | 16/11         | 4/16          | 12/4          |
| Age (years)        | 34 (19–65)    | 43 (24–78)    | 43 (27–54)    |
| BMI (kg·m⁻²)       | 24.6±3.4      | 25.5±4.3      | 27.4±5.6      |

**Lung function**

- FEV₁ (% predicted) 105.0±14
- FVC (% predicted) 105.0±15
- FEV₁/FVC (%) 83±6.0
- TLC (% predicted) 101±23
- FRC̄pleth (% predicted) 97.0±27
- DLCO (% predicted) 102±13
- FRC̄MBNW (L)³ 2.94±0.89
- LCI (TO) ² 6.49±0.47
- LCI (TO)# 6.49±0.47
- S̄cond (L⁻¹)⁴ 0.019±0.011
- S̄acin (L⁻¹)⁵ 0.056±0.020

Data are presented as mean±SD or median (interquartile range) unless otherwise stated. BMI: body mass index; DLCO: diffusing capacity of the lung for carbon monoxide; FEV₁: forced expiratory volume in 1 s; FVC: functional residual capacity; FRC: functional residual capacity; LCI: lung clearance index; MBNW: multiple breath nitrogen washout; S̄acin: distal/intra-acinar airs ventilation heterogeneity; S̄cond: conducting airways ventilation heterogeneity; TLC: total lung capacity; TO: lung turnover. *Corrected, controlled breathing protocol values used. †p<0.05 health versus asthma. ‡p<0.05 health versus smokers. ¶p<0.05 asthma versus smokers.

Compared to FRC̄pleth in all three groups (mean±SD differences of −0.26±0.47 L (p=0.008), −0.26±0.37 L (p=0.006) and −0.64±0.71 L (p=0.003) in health, asthma and smokers, respectively).

Notably, mean (95% CI) S̄cond significantly increased by 11.1 (−1.4–23.5)%, 14.0 (4.2–23.9) and 36 (19.8–52.2)% following sensor correction in health, asthma and smokers, respectively. In contrast, S̄acin was significantly lower following sensor correction in the asthma and smokers groups, with a trend to significance in the healthy group (p=0.08). The impact on S̄acin, however, was minimal with mean decreases (95% CI) of 1.8 (0.4–4.0)%, 2.9 (0.9–4.9) and 4.8 (0.7–8.9)% observed in health, asthma and smokers.

TABLE 2  Effects of sensor correction on main MBNW parameters

| Parameter          | Standard | Corrected | Mean difference (95% CI) absolute | Mean difference (95% CI) relative | p-value |
|--------------------|----------|-----------|-----------------------------------|----------------------------------|---------|
| **Health**         |          |           |                                   |                                  |         |
| FRC (L)            | 3.19±0.98| 2.94±0.89 | −0.25 (−0.29–−0.21) | −7.7 (−8.4–−7.0) | <0.0001 |
| LCI (TO)           | 7.20±0.58| 6.49±0.47 | −0.71 (−0.80–−0.63) | −9.8 (−10.8–−8.8) | <0.0001 |
| S̄cond (L⁻¹)       | 0.017±0.009| 0.019±0.011| 0.002 (0.003–0.004) | 0.011 (1.4–23.5) | 0.03    |
| S̄acin (L⁻¹)       | 0.057±0.020| 0.056±0.020| 0.0009 (−0.002–0.0001) | −1.8 (−4.0–−0.44) | 0.08    |
| **Asthma**         |          |           |                                   |                                  |         |
| FRC (L)            | 2.84±0.76| 2.62±0.72 | −0.22 (−0.26–−0.18) | −8.0 (−9.2–−6.8) | <0.0001 |
| LCI (TO)           | 7.94±1.19| 7.23±1.04 | −0.71 (−0.84–−0.58) | −8.8 (−10.0–−7.6) | <0.0001 |
| S̄cond (L⁻¹)       | 0.029±0.016| 0.033±0.018| 0.004 (0.001–0.006) | 0.014 (4.2–23.9) | 0.003   |
| S̄acin (L⁻¹)       | 0.088±0.049| 0.086±0.049| −0.002 (−0.003–0.0008) | −2.9 (−4.9–−0.9) | 0.003   |
| **Smokers**        |          |           |                                   |                                  |         |
| FRC (L)            | 3.19±0.91| 2.93±0.84 | −0.28 (−0.34–0.22) | −8.4 (−10.2–−6.7) | <0.0001 |
| LCI (TO)           | 7.69±0.96| 7.05±0.82 | −0.65 (−0.78–−0.51) | −9.3 (−9.7–−6.8) | <0.0001 |
| S̄cond (L⁻¹)       | 0.029±0.015| 0.038±0.019| 0.009 (0.005–0.013) | 0.036 (19.8–52.2) | 0.0004  |
| S̄acin (L⁻¹)       | 0.10±0.040| 0.095±0.037| −0.004 (−0.007–0.001) | −4.8 (−8.9–−0.71) | 0.008   |

Effects of sensor correction on main MBNW parameters. Data are presented as mean±SD unless otherwise stated. FRC: functional residual capacity; LCI: lung clearance index; MBNW: multiple breath nitrogen washout; S̄acin: distal/intra-acinar airs ventilation heterogeneity; S̄cond: conducting airs ventilation heterogeneity; TO: lung turnover. The standard (uncorrected) value is the reference. Absolute difference is calculated as corrected – standard. Relative difference is calculated as corrected – standard/standard × 100.
FIGURE 1  Comparison of standard (uncorrected) and corrected multiple breath nitrogen washout (MBNW) parameters in health. There were strong correlations between uncorrected and corrected functional residual capacity (FRC) (r=0.99 and p<0.0001) (a), lung clearance index (LCI) (r=0.94 and p<0.0001) (b), $S_{\text{cond}}$ (r=0.86, p<0.0001) (c) and $S_{\text{acin}}$ (r=0.98, p<0.0001) (d). Bland-Altman plots show that sensor correction results in a lower FRC (mean difference (95% limits of agreement) –0.25 (–0.46–0.04), p<0.0001) (e), lower LCI (–0.71 (–1.14–0.28), p<0.0001) (f), higher $S_{\text{cond}}$ (0.002 (–0.007–0.012), p=0.027) (g), but no change in $S_{\text{acin}}$ (–0.009 (–0.006–0.004), p=0.08) (h). There was also significant proportional bias confirmed by linear regression for FRC (p<0.0001), LCI (p=0.007) and $S_{\text{cond}}$ (p<0.003)(h). There was also significant proportional bias confirmed by linear regression for FRC (p=0.02) and LCI (p=0.01). $S_{\text{acin}}$: distal/intra-acinar airways ventilation heterogeneity; $S_{\text{cond}}$: conducting airways ventilation heterogeneity.

FIGURE 2  Comparison of standard (uncorrected) and corrected multiple breath nitrogen washout (MBNW) parameters in asthma. There were strong correlations between uncorrected and corrected functional residual capacity (FRC) (r=0.99 and p<0.0001) (a), lung clearance index (LCI) (r=0.94 and p<0.0001) (b), $S_{\text{cond}}$ (r=0.94, p<0.0001) (c) and $S_{\text{acin}}$ (r=0.98, p<0.0001) (d). Bland-Altman plots show that sensor correction results in a lower FRC (mean difference (95% limits of agreement) –0.22 (–0.38–0.06), p<0.0001) (e), lower LCI (–0.71 (–1.24–0.18), p<0.0001) (f), higher $S_{\text{cond}}$ (0.004 (–0.006–0.014), p=0.003) (g) and lower $S_{\text{acin}}$ (–0.002 (–0.007–0.003), p=0.003) (h). There was also significant proportional bias confirmed by linear regression for FRC (p=0.02) and LCI (p=0.01). $S_{\text{acin}}$: distal/intra-acinar airways ventilation heterogeneity; $S_{\text{cond}}$: conducting airways ventilation heterogeneity.

https://doi.org/10.1183/23120541.00614-2021
smokers, respectively. When using the free-breathing protocol, similar effects for LCI, FRC, $S_{\text{cond}}$ and $S_{\text{acin}}$ were observed in health and asthma (Online Supplement, Table S1).

There were strong correlations between all corrected and uncorrected MBNW values across the three groups (all r-values >0.85) (figures 1–3, panels A–D) and for both breathing protocols (Online Supplement, Figures S1 and S2). Bland–Altman plots showed that the effect of sensor correction on LCI and FRC demonstrated strong proportional bias in all three groups (greater difference with higher mean value) (figures 1–3, panels E–H). The Bland–Altman plots also revealed large variance in $S_{\text{cond}}$ and significant proportional bias in health and smokers, but not in asthma. Less variance in differences was seen in $S_{\text{acin}}$ and there was no evidence of proportional bias in any of the three groups.

**Effects on within- and between-session repeatability in health**

Fifteen healthy volunteers underwent repeat testing. Within-session and between-session variability measurements are presented in table 3. There were no differences observed in within-session CoVs between corrected and uncorrected FRC (p=0.46) or LCI (p=0.84). Between-session variability was minimally affected by the sensor error. Corrected FRC and LCI showed narrower 95% LoAs, whereas $S_{\text{acin}}$ and $S_{\text{cond}}$ showed slightly wider 95% LoAs. Between-session ICC values were numerically comparable between corrected and uncorrected values. Similar impact on within- and between-session repeatability was observed with the free-breathing protocol (Online Supplement, Table S2).

**Discussion**

In this study, we demonstrate that correction of the O₂ and CO₂ sensor error in the Exhalyzer D system results in significantly lower FRC and LCI, and higher $S_{\text{cond}}$ values in three different adult patient groups. The impact on $S_{\text{acin}}$, although statistically significant, was minimal. There were strong correlations between the corrected and uncorrected values for all MBNW parameters in all three groups. Importantly, the effect of the correction showed a significant proportional bias in FRC and LCI in all three groups, and significant proportional bias in $S_{\text{cond}}$ was also evident in health and smokers, although not in asthma. The O₂ and CO₂ sensor error correction produced less variance in $S_{\text{acin}}$ compared to other parameters and there was no evidence of proportional bias. Furthermore, sensor error correction had minimal impact on within-session and between-session variability, with a smaller 95% LoA for LCI between sessions.

**FIGURE 3** Comparison of standard (uncorrected) and corrected multiple breath nitrogen washout (MBNW) parameters in smokers. There were strong correlations between uncorrected and corrected functional residual capacity (FRC) ($r=0.99$ and $p<0.0001$) (a), lung clearance index (LCI) ($r=0.97$ and $p=0.0001$) (b), $S_{\text{cond}}$ ($r=0.93$, $p=0.0001$) (c) and $S_{\text{acin}}$ ($r=0.99$, $p=0.0001$) (d). Bland–Altman plots show that sensor correction results in a lower FRC (mean difference (95% limits of agreement) $–0.28$ ($–0.50$–$–0.06$), $p=0.0001$) (e), lower LCI ($–0.65$ ($–1.16$–$–0.13$), $p=0.0001$) (f), higher $S_{\text{cond}}$ ($0.009$ ($–0.006$–$0.025$), $p=0.0004$) (g) and lower $S_{\text{acin}}$ ($–0.04$ ($–0.015$–$–0.006$), $p=0.008$) (h). There was also significant proportional bias confirmed by linear regression for FRC ($p=0.04$), LCI ($p=0.03$) and $S_{\text{cond}}$ ($p=0.01$). The uncorrected value is the reference. Absolute differences were calculated as corrected–uncorrected. $S_{\text{acin}}$: distal/intra-acinar airways ventilation heterogeneity; $S_{\text{cond}}$: conducting airways ventilation heterogeneity.
Overestimation of FRC and LCI by the Exhalyzer system was first suggested when comparing the use of sulfur hexafluoride (SF₆) to N₂ as a tracer gas. Jensen et al. [21] found in children with CF that N₂ resulted in higher estimates of FRC and LCI compared to SF₆ obtained using mass spectrometry. In addition to differences in the diffusion front, the assumption was that back-secretion of N₂ from the tissues probably contributed to overestimation of FRC by MBNW. In fact, subsequent device comparison studies in adults tended to show FRC by the Exhalyzer D system to be larger than FRCpleth [22, 23]. However, these findings are at odds with the idea that gas dilution techniques during tidal breathing can only access communicating lung units and not trapped gas compartments, such that the estimated FRC in disease should be lower than FRC obtained from plethysmography, which includes all compressible gas volume within the lungs. Prior to reanalysis, there were no differences between FRCpleth and FRCMBNW in smokers, patients with asthma or in health but sensor error correction resulted in a significantly lower FRCMBNW compared to FRCpleth in all groups, more consistent with expectation. These results suggest that the sensor error explains most of the overestimation of FRC seen in the Exhalyzer device, just as Sandvik et al. [13] found that sensor error correction of MBNW removed the discrepancy in FRC between N₂ and SF₆. It is unknown whether the error affects different commercially available MBW utilising O₂ and/or CO₂ sensors, a subject that warrants further investigation.

Our study is the first to demonstrate the impact of the O₂ and CO₂ sensor error correction on FRC and LCI in adults, and the first to investigate the impact on Scond and Sacin. The effect of the sensor error on LCI and FRC has been described previously in infants and children, and our data are consistent with their findings in both magnitude and presence of proportional bias [12, 13]. The alignment of these findings is important to understand consistency in the correction algorithm. The high correlations between uncorrected and corrected values suggest that previous findings involving correlations with MBNW indices may be preserved, but the presence of significant proportional bias indicates that previous studies examining interventional effects will require reanalysis, both to reconfirm previous findings and to allow comparability with future studies. Although a recent reanalysis of CF clinical trials was reassuring to a degree and showed that while treatment effects were reduced, they were maintained following sensor correction [14].

Previous studies investigating the effect of sensor error correction were in infants and children [12–14], hence they did not include a comparison of phase III slope indices Scond and Sacin, which are not as commonly used in paediatric compared to adult age groups. Scond is calculated as the slope of the plot of normalised phase III slope (SnIII) versus lung turnover (TO), between TO 1.5 and 6, where SnIII is the slope of phase III in the N₂ expirogram normalised by mean or end-tidal N₂ concentration. Errors in Scond arise from two sources. First, the observed overestimation of FRC results in a lower TO, shortening the SnIII versus TO plot leftward and slightly elevating Scond. Second, as the washout progresses towards higher values of TO, the phase III slope is normalised by an increasingly overestimated N₂ concentration. The effect is a less steep SnIII versus TO plot, thus lowering calculated Scond. These effects are demonstrated in figure 4, where corrected SnIII values for three different patients are increased, resulting in larger Scond as calculated between TO 1.5 and 6. In particular, the dominant effect of the impact on SnIII...
is clearly seen in panel 4C where uncorrected SnIII values deviate markedly from the corrected values at high TO. However, the change in SnIII in the first breath was minimal, both because the sensor error is smallest at high N₂ concentrations, and because the N₂ concentration used for normalisation is large at this point in the washout. Much of the effect of sensor correction on $S_{\text{acin}}$ probably comes from propagation of the $S_{\text{cond}}$ error into the correction applied to SnIII(1) to obtain $S_{\text{acin}}$ [20].

Our comparison found $S_{\text{cond}}$ to be significantly increased by the sensor error correction, and furthermore with a significant proportional bias in both health and in smokers. However, this distinction between groups is probably a manifestation of small numbers in each cohort, coupled with the inherent variability in the measurement of $S_{\text{cond}}$. Indeed, when the three cohorts are combined into the single dataset (Online Supplement, Figure S3), it is clear that the sensor effort correction results in comparable effects on $S_{\text{cond}}$ regardless of the underlying pathophysiology.

Correction of the sensor error resulted in minimal impact on within-session and between-session variability in health. Within-session CoV remained small in FRC and LCI, demonstrating that trial repeatability for MBNW was high even after reanalysis. Similarly, all parameters had minimal change in between-session difference, with a small change in the LoA for LCI, which is probably attributed to the overall reduction in LCI caused by correction. Furthermore, we also reanalysed previously published data collected using both free breathing and controlled breathing [15, 16]. Sensor error correction did not affect the between-protocol differences in $S_{\text{cond}}$ and $S_{\text{acin}}$ in health [15] or asthma [16], nor their dependences on the breathing pattern.

This study is limited by the selection criteria for the previous studies that we have included for reanalysis. Patients with asthma had relatively mild disease, and smokers were recruited for a larger study based on having abnormal ventilation heterogeneity as described in the Methods, and thus may not be representative of the population in general. Future reanalysis of MBNW data is required to understand the effect of sensor error correction in disease more broadly and the associated implications. Moreover, in our reanalysis, we chose to retain the same breath exclusions and other settings in the original analysis, to allow us to solely examine the effect of corrected N₂ concentrations on MBNW indices. There is a chance that the adjusted washout traces may result in, for example, changes in the shape of the expirogram, which may result in different quality control decisions by a manual operator. However, we attempted to maintain a consistent approach for quality control. The new software version also includes changes in the way in which delay between flow and gas concentration sampling is calculated, to include a dynamic delay correction [24], which was not implemented in our reanalysis, but which may be a factor affecting comparability between old and new studies in the literature involving the Exhalyzer D. This was intentionally done to focus on the effects of the cross-talk sensor error correction.

In conclusion, our study is the first to describe the effect of O₂ and CO₂ sensor error correction on the Exhalyzer D MBNW system in adults, and the first to investigate the effect on $S_{\text{cond}}$ and $S_{\text{acin}}$. Our results confirm the LCI and FRC effects seen in infants and children and demonstrate strong underestimation with proportional bias for $S_{\text{cond}}$, with errors up to 50% observed in those with the greatest ventilation heterogeneity, but minimal effects on $S_{\text{acin}}$. While the discovery of the error is an important step towards
improved accuracy of MBNW devices, it also represents an important hurdle for ongoing efforts to support MBNW as a clinical tool or an end-point for clinical studies. These findings provide important considerations for the interpretation of previously published adult MBNW studies, and those in younger age groups incorporating phase III slope analysis. The magnitude of effect supports reanalysis of that data to better understand the true findings.

Provenance: Submitted article, peer reviewed.

Acknowledgements: We would like to acknowledge the study participants for volunteering the time and effort required to conduct this study. We would also like to thank Blake Handley (Dept of Respiratory Medicine, Royal North Shore Hospital, St Leonards, NSW, Australia) and Stephen Milne (Centre for Heart Lung Innovation and Division of Respiratory Medicine, University of British Colombia, Vancouver, Canada) for assistance with the healthy and asthma datasets, and Prof. Christine Jenkins (ECOS Study; Dept of Thoracic Medicine, Concord Hospital, Concord, Australia and The George Institute for Global Health, Sydney, Australia) for provision of the smokers dataset used for this study.

Data sharing statement: The study protocol and raw data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of interest: No conflicts of interest, financial or otherwise, relating to this study are declared by the authors.

Support statement: S. Rutting was supported by the Berg Family Foundation. The smokers dataset was from a larger study funded by an investigator-initiated grant from GlaxoSmithKline, Australia. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1. Verbanck S, Schuermans D, Vincken W. Inflammation and airway function in the lung periphery of patients with stable asthma. *J Allergy Clin Immunol* 2010; 125: 611–616.
2. Verbanck S, Schuermans D, Van Muylem A, et al. Conductive and acinar lung-zone contributions to ventilation inhomogeneity in COPD. *Am J Respir Crit Care Med* 1998; 157: 1573–1577.
3. Zimmermann SC, Tonga KO, Thamrin C. Dismantling airway disease with the use of new pulmonary function indices. *Eur Respir Rev* 2019; 28: 180122.
4. Ratjen F, Davis SD, Stanojevic S, et al. Inhaled hypertonic saline in preschool children with cystic fibrosis (SHIP): a multicentre, randomised, double-blind, placebo-controlled trial. *Lancet Respir Med* 2019; 7: 802–809.
5. Ratjen F, Hug C, Marigowda G, et al. Efficacy and safety of lumacatror and ivacaftor in patients aged 6–11 years with cystic fibrosis homozygous for F508del-CFTR: a randomised, placebo-controlled phase 3 trial. *Lancet Respir Med* 2017; 5: 557–567.
6. Farah CS, King GG, Brown NJ, et al. Ventilation heterogeneity predicts asthma control in adults following inhaled corticosteroid dose titration. *J Allergy Clin Immunol* 2012; 130: 61–68.
7. Farah CS, King GG, Brown NJ, et al. The role of the small airways in the clinical expression of asthma in adults. *J Allergy Clin Immunol* 2012; 129: 381–387, 387.e1.
8. Tang FSM, Rutting S, Farrow CE, et al. Ventilation heterogeneity and oscillometry predict asthma control improvement following step-up therapy in uncontrolled asthma. *Respirology* 2020; 25: 827–835.
9. Farah CS, Badal T, Reed N, et al. Mepolizumab improves small airway function in severe eosinophilic asthma. *Respir Med* 2019; 148: 49–53.
10. Jetmalani K, Thamrin C, Farah CS, et al. Peripheral airway dysfunction and relationship with symptoms in smokers with preserved spirometry. *Respirology* 2018; 23: 512–518.
11. Jetmalani K, Chapman DG, Thamrin C, et al. Bronchodilator responsiveness of peripheral airways in smokers with normal spirometry. *Respirology* 2016; 21: 1270–1276.
12. Wyler F, Oestreich MH, Frauchiger BS, et al. Correction of sensor crosstalk error in Exhalyzer D multiple-breath washout device significantly impacts outcomes in children with cystic fibrosis. *J Appl Physiol (1985)* 2021; 131: 1148–1156.
13. Sandvik RM, Gustafsson PM, Lindblad A, et al. Improved agreement between N(2) and SF(6) multiple-breath washout in healthy infants and toddlers with improved EXHALYZER D sensor performance. *J Appl Physiol (1985)* 2021; 131: 107–118.
14. Robinson PD, Jensen R, Seeto RA, et al. Impact of cross-sensitivity error correction on representative nitrogen-based multiple breath washout data from clinical trials. *J Cyst Fibros* 2021; 21: e204–e207.
15. Handley BM, Jeagal E, Schoefell RE, et al. Controlled versus free breathing for multiple breath nitrogen washout in healthy adults. *ERJ Open Res* 2021; 7: 00435-2020.
16 Handley BM, Bozier J, Jeagal E, et al. Controlled versus free breathing for multiple breath nitrogen washout in asthma. *ERJ Open Res* 2021: 00487-2021.

17 Quanjer PH, Stanojevic S, Cole TJ, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.

18 Quanjer PH, Tammeling GJ, Cotes JE, et al. Lung volumes and forced ventilatory flows. *Eur Respir J* 1993; 6: Suppl. 16, 5–40.

19 Verbanck S, Schuermans D, Paiva M, et al. Mitigating increased variability of multiple breath washout indices due to tidal breathing. *Eur Respir J* 2021; 57: 2002765.

20 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.

21 Jensen R, Stanojevic S, Gibney K, et al. Multiple breath nitrogen washout: a feasible alternative to mass spectrometry. *PLoS One* 2013; 8: e56868.

22 Poncin W, Singer F, Aubriot AS, et al. Agreement between multiple-breath nitrogen washout systems in children and adults. *J Cyst Fibros* 2017; 16: 258–266.

23 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.

24 Gustafsson PM, Bengtsson L, Lindblad A, et al. The effect of inert gas choice on multiple breath washout in healthy infants: differences in lung function outcomes and breathing pattern. *J Appl Physiol (1985)* 2017; 123: 1545–1554.