

Online supplement

Multiple-breath washout measurements can be significantly shortened in children

Sophie Yammine*, Florian Singer*, Chiara Abbas, Markus Roos, Philipp Latzin

*Sophie Yammine and Florian Singer contributed equally to this work

Division of Respiratory Medicine, Department of Paediatrics, University Children's Hospital of Bern, Switzerland.

Corresponding author: Philipp Latzin, MD, PhD; Division of Respiratory Medicine, Department of Paediatrics, University Children's Hospital of Bern, 3010 Bern, Switzerland;
Tel: +41 31 632 9437; Fax: +41 31 632 4807; Email: philipp.latzin@insel.ch

Methods

Study subjects

In this prospective cross-sectional study 50 children with cystic fibrosis (CF) and 30 healthy controls aged between six and seventeen years at the Children's University Hospital of Bern were enrolled between November 2011 and May 2012. Children with CF performed nitrogen multiple-breath washout (N₂MBW) and spirometry in that order, controls performed N₂MBW. Outcomes were lung clearance index (LCI), functional residual capacity (FRC), and forced expiratory volume in one second (FEV₁) reported in z-scores (1). Data from these children have not been published before. We recruited children with confirmed CF attending the outpatient clinic and age-matched healthy controls from local schools and playgroups. Exclusion criteria for children with CF were fever or acute respiratory deterioration. Exclusion criteria for controls were doctor diagnosis of asthma or history of respiratory tract infection within the last four weeks. All children performed three N₂MBW according to current guidelines (2).

Nitrogen multiple-breath washout

Three complete N₂MBW until 1/40th of the starting N₂ end-tidal concentration (cet), *i.e.* 2.5% N₂ cet from initial normalized N₂ cet, were performed as described previously (3;4). We used a recently validated setup (Exhalyzer D and Spiroware 3, Eco Medics AG, Duernten, Switzerland) (3-5). Children performed N₂MBW while sitting upright, wearing a nose clip, and tidally breathing through a snorkel mouth piece. N₂MBW was stopped after three breaths below 2.5% N₂ cet. Time for measurements and between tests (measurement time plus one minute) was recorded. End-tidal N₂ cet was controlled prior to the next N₂MBW. N₂MBW runs were accepted if FRC varied less than 20% (6).

Early stop in multiple-breath washout

We calculated standard LCI (LCI_{std}) at the classical time point, *i.e.* $1/40^{th}$ of the starting N_2 conc. = 2.5% N_2 cet, and nine earlier alternative time-points of the three N_2 MBW curves from each child. Using customized software (Matlab R2006a; The MathWorks, Inc.) we calculated the standard LCI (LCI_{std}) at 2.5% N_2 cet and nine LCI from empirically chosen N_2 cet cut-offs: 18% (LCI_{18}), 15% (LCI_{15}), 12% (LCI_{12}), 9% (LCI_9), 7% (LCI_7), 6% (LCI_6), 5% (LCI_5), 4% (LCI_4), 3% (LCI_3), and the standard 2.5% ($LCI_{2.5} = LCI_{std}$) N_2 cet. All LCI were calculated as cumulative expired volume/FRC at the respective cut-offs. Standard and alternative LCI give the number of lung turn overs required to washout out the lungs until the respective target N_2 cet is reached, thus respective LCI units are lung turnovers (TO).

Less than three multiple-breath washout tests

From standard N_2 MBW until 2.5% N_2 cet, we assessed duration and diagnostic performance of (i) single LCI (LCI_{single}) of the first N_2 MBW run and of (ii) mean LCI (LCI_{double}) of the first and second N_2 MBW runs in comparison to the standard mean LCI (LCI_{std}) of three runs.

Ethics statement

The study was approved by the Ethics Committee of the Canton of Bern, Switzerland. The children`s assent was obtained and parents or caregivers provided written informed consent for this study.

Statistical Methods

A recruitment goal of 80 children was chosen to match our sample size estimation ($N = 60$) taking into account a failure rate of approximately 20% as described previously in school-aged children (7-9). Prior to analysis we defined quality criteria to be met by alternative LCI measurements: (i) Predictive value (R^2) for $LCI_{std} > 90\%$ or limits of agreement $\leq 5\%$ (ii)

coefficient of variation (CV) or relative difference between two LCI $\leq 7\%$ and (iii) sensitivity and specificity within $\pm 3\%$ of LCI_{std}. R^2 for LCI_{std} was calculated from uni-variable linear regression models. Intra-test repeatability of LCI from three N₂MBW runs (LCI_{std}, LCI₁₈₋₃) was calculated as intra-test CV. Intra-test repeatability of LCI_{double} was calculated as the relative difference (%) of the first minus the second LCI divided by the mean of both LCI (LCI_{double}). Agreement between LCI_{single} and LCI_{double} with LCI_{std} was assessed by Bland Altman plots (10). By means of receiver operating characteristic curves (ROC) and LCI cross-plots, concordance between LCI, and sensitivity and specificity for CF lung disease were assessed. Upper limit of normal (ULN; mean LCI + 1.96 standard deviations [SD]) and LCI z-scores were defined from the healthy children. LCI z-score values were calculated from the corresponding LCI values in a child with CF minus the mean LCI of the healthy children divided by the LCI SD of healthy children. Thus normal ranges for standard and alternative LCI were based on the study's healthy cohort. Test duration was compared by paired t-test. P-values < 0.05 were considered statistically significant. All analyses were done using StataTM (Stata Statistical Software: Release 11. College Station, TX: StataCorp LP).

Results

Technically acceptable triplicate N₂MBW were obtained in 68 (31 boys) children (44 with CF). Reasons for exclusion were leaks in 7 and FRC variability > 20% in 5 children. Mean (range) age of children with CF was 12.0 (5.7 to 16.9) and 12.4 (7.3 to 15.9) years in controls. ULN for LCI_{std} was 7.56 (6.44+[1.96*0.57]) (Table 1). LCI_{std} identified 36 (81.8%) CF children with abnormal lung function (LCI_{std}> 7.56 TO and >2 z-scores). 27 (61.4%) children with CF had normal FEV₁ z-scores (< -2 z-scores) but abnormal LCI_{std} (> 2 z-scores). Mean (SD) FRC intra-test CV in CF was 6.2% (4.2%) and 5.4% (3.7%) in controls. FRC variability was < 10% in 56 children (82%), better than 15% in 11 children (16%), and 17.5% in one child.

Early stop in multiple-breath washout

Calculating standard and shortened LCI was feasible in all N₂MBW measurements (n = 204, Table 1). LCI₅, reflecting N₂ washout until 1/20th, was the earliest time point in N₂MBW to meet the predefined quality criteria (i) R² > 90%, (ii) CV < 7%, and (iii) sensitivity and specificity within 3% of LCI_{std} (Table 1 and 2). LCI₅ predicted 93% of LCI_{std} (Figure 1). CV of LCI_{5%} was 5.1%, compared to 5.4% of LCI_{std} (Table 2). LCI₅ and LCI_{std} were both equally specific (100%) and provided a sensitivity of 82.6% and 83.0% for CF lung disease. ULN for LCI₅ was 5.53 (4.86+[1.96*0.34]). LCI₅ identified 35 (79.5%) CF children with abnormal lung function (LCI₅> 5.53 TO and >2 z-scores). LCI₅ and LCI_{std} concordantly classified 63 (92.6%) children (Figure 1). Mean (range) test duration for LCI₅ in CF was 8.8 (3.8-18.4) min. and in controls 8.5 (5.0-12.1) min. being significantly shorter than for LCI_{std} in CF 12.6 (4.7-26) min. and controls 10.8 (5.9-15.6) min.. This represents an average time saving of 30% in CF (p < 0.001) and 21% in controls (p < 0.001, Table 1).

Every LCI calculated at cut-offs earlier than 5% N₂ cet was highly associated with LCI_{std} and had low intra-test variability (CV) ranging from 4.8% to 5.4% in CF and from 5.0% to 6.8%

in controls. However, all time points for LCI earlier than 5% N₂ cet provided less sensitivity and specificity with more than 3% difference compared to LCI_{std} (Table 1).

Less than three multiple-breath washout tests

LCI_{single} and LCI_{double} were very similar to and highly associated with LCI_{std}, and both had similarly high diagnostic performance for CF lung disease (Table 3). Prediction of LCI_{std} (R²) was 97% and 99%, respectively (Figures 2 and 3). Sensitivity and specificity were similar or even higher compared to LCI_{std} (Table 3). LCI_{single} and LCI_{double} were significantly faster obtained than LCI_{std}. For LCI_{double}, relative mean (SD) intra-test difference between the first and second run was 6.8 (4.9)%. Mean (range) test duration for LCI_{single} was 2.2 (0.4-5.2) min. in CF and 1.7 (0.9-2.8) min. in healthy controls, for LCI_{double} 7.4 (2.4-16.0) min. in CF and 6.2 (3.5-9.4) min. in healthy controls (Table 3). Average time savings were 82% in CF and 84% in controls for LCI_{single}, and 41% in CF and 42% in controls for LCI_{double} (all p < 0.001). However, only LCI_{double} was within 5% limits of agreement compared to LCI_{std} (data not shown).

Based on these data, the shortest protocol for N₂MBW would be LCI₅ from the first two N₂MBW runs. However, LCI₅ based on two runs predicted only 88% of LCI_{std}.

Discussion

Comparison to previous studies

N₂MBW can be reliably and significantly abbreviated. LCI₅ and LCI_{double} predicted 93-99% of LCI_{std} and provided the same or better diagnostic performance for pediatric CF lung disease. These data may help implementing LCI measurement in clinical routine. Previous studies focused on classical tracer gas MBW until 1/40th showing reliability of LCI across age groups (7;11) or LCI sensitivity for *e.g.* structural changes in chest CT scans (12;13). None of these studies primarily aimed to increase practicability of LCI measurement for clinical routine. Recently one retrospective study in a large pediatric cohort (n = 198) elegantly showed that variability in FRC between two tests may approach 27% in children with CF and controls (6). Despite this variability, LCI derived from two sulfur hexafluoride (SF₆)MBW runs classified 94% of children correctly. While these data confirm our findings (concordance of LCI_{double} = 100%, Figure 3), they may be less applicable for clinical routine. Hard- and software used by Robinson *et al.* (6) are not available any longer and SF₆ use is restricted in many countries. The setup did not allow for real-time monitoring of FRC and LCI which certainly is important to prospectively assess whether two or three MBW are needed.

Abbreviated N₂MBW protocols in our study significantly reduced test duration. These protocols potentially increase feasibility in routine care and acceptance in younger children. Previous studies have reported MBW success rates of 80% on average in school-age children for three valid MBW runs (7-9). Possible reasons for test failure were leaks or coughing and irregular breathing due to increasing discomfort. Earlier stop of MBW will *per se* increase success rates in any setting.

LCI_{double} seems especially attractive in young children with predominantly mild CF lung disease. Complete N₂MBW until 1/40th is usually achieved fast in mild CF lung disease. The LCI_{double} provided the highest sensitivity in our study and thus provides sufficient confidence

to detect mild CF lung disease. The LCI_5 may be more applicable *e.g.* in advanced lung disease where N_2 MBW requires much time, and LCI measurement is more variable within test sessions. Above an LCI_{std} of 10 correlation with LCI_5 gets poorer (Figure 1), mostly with a slight underestimation of LCI_{std} by LCI_5 . The trend of higher LCI_{std} compared to LCI_5 may be due to extremely slowly ventilated lung regions at the end of long washout and possible bias from tissue N_2 for which no standard correction algorithm exists (14). Bias from tissue N_2 may decrease with abbreviated N_2 MBW test duration, thus LCI_5 may be even more accurate in advanced lung disease than LCI_{std} . Which MBW protocols are best used in respective patient groups need to be established in future as well as comparison of long-term repeatability and response to intervention between MBW protocols.

Physiological aspects

We are the first to show that N_2 MBW is not required until $1/40^{th}$ of initial N_2 cet to measure LCI. Our data clearly indicate that during the last third of N_2 MBW, sensitivity and specificity of LCI only marginally increase while test duration significantly increases. Given the exponential decrease of N_2 cet as also previously described (15), it is not surprising that LCI measured at N_2 cet targets in the flatter tail of the N_2 MBW curve contain very similar information. The shape of the N_2 MBW curve also suggests at least two functional lung compartments being sequentially washed out with the second phase believed to reflect peripheral airways (16). This portioned washout behavior seems even more pronounced in advanced CF lung disease with prolonged washout in slowly ventilated lung regions. This probably explains lower agreement between LCI_5 and LCI_{std} in those subjects with high LCI_{std} values (Figure 1). Nevertheless, we found that N_2 MBW until $1/20^{th}$ is sufficiently robust in the current study population consisting mainly of mild CF lung disease regarding the predefined quality measures. Despite lower agreement between LCI_5 and LCI_{std} for higher

LCI_{std} values, LCI₅ might be an even more robust marker of ventilation inhomogeneity in advanced CF lung disease, being less susceptible to bias from tissue N₂.

We also show that performance of LCI_{double} measured from two N₂MBW is very similar to LCI_{std} from three N₂MBW. To our knowledge the only reason for three tests in previous studies is that tidal breathing online quality control is scarce thus requiring *post-hoc* quality control, *i.e.* calculating CV. Indeed, full FRC stability within tests cannot be achieved in children and even in adults (3;4). However, LCI is more robust than FRC. LCI, a volume ratio, partly cancels out variability of the denominator FRC. Our and previous data suggest that LCI from two MBW provide sufficient sensitivity and specificity for pediatric CF lung disease. This is even more true as modern MBW setups display flow-volume-loops, CEV, and FRC for online quality control. Considering potential bias on airway calibers due to prolonged dry tracer gas exposure, less than three MBW tests seem definitely preferable.

Methodological considerations

LCI₅ and LCI_{double} reliably identified altered gas mixing efficiency in children with CF. The N₂ cut-offs and quality criteria were defined prior to analyses. However, cut-offs and quality criteria are arbitrary to some extent and further studies assessing these in different CF populations and other disease groups such as primary ciliary dyskinesia or sickle cell disease seem important (17). Whether our findings apply to customized equipment and other tracer gases needs to be determined.

Similar to most other MBW studies we used “CF diagnosis” as proxy for assessing sensitivity and specificity of LCI. Previous data suggest reasonable agreement between actual structural airway pathology detected in HRCT and LCI (12;13;18) . In our study, 36 (82%) children with CF had abnormal LCI_{std}, which is in agreement with the prevalence in other pediatric CF cohorts (12). However, we did not apply HRCT to ascertain abnormal LCI findings. Further,

the impact of functional or radiological measures of small airway pathology on CF care and respiratory disease outcomes remains to be determined.

Additional analysis of N₂MBW, *i.e.* phase III slopes for Scond and Sacin measurement, may be feasible from abbreviated N₂MBW protocols. Due to the inherent variability of Scond and Sacin, two tests may be insufficient (3). However, Scond and Sacin were not the focus of this paper and may be less suitable for routine assessments (19;20).

FIGURE LEGENDS

Figure 1. Early stop in multiple-breath washout. Concordance of lung clearance index (LCI) measured at 5% end-tidal nitrogen concentration (N_2 cet) vs. standard LCI at 2.5% N_2 cet from three N_2 MBW. Healthy children ($n = 24$) are displayed as closed circles and children with CF ($n = 44$) as open circles. The dashed lines correspond to the upper limits of normal LCI (mean LCI + 1.96 standard deviations from healthy controls).

Figure 2. Less than three multiple-breath washout tests - single LCI. Concordance of lung clearance index (LCI) from the first nitrogen multiple-breath washout run vs. standard LCI (mean of three runs). Healthy children ($n = 24$) are displayed as closed circles and children with CF ($n = 44$) as open circles. The solid black line gives the line of identity. The dashed lines correspond to the upper limits of normal LCI (mean LCI + 1.96 standard deviations from healthy controls).

Figure 3. Less than three multiple-breath washout tests - double LCI. Concordance of lung clearance index (LCI) from the first and second nitrogen multiple-breath washout runs vs. standard LCI (mean of three runs). Healthy children ($n = 24$) are displayed as closed circles and children with CF ($n = 44$) as open circles. The solid black line gives the line of identity. The dashed lines correspond to the upper limits of normal LCI (mean LCI + 1.96 standard deviations from healthy controls).

Table 1: The early stop - standard and alternative lung clearance index at different N₂ concentrations.

| N ₂ cet (%) | ULN | Time (min) | | Sens. (%) | Spec. (%) | Correct (%) | R ² |
|------------------------|------|------------|-----------|-----------|-----------|-------------|----------------|
| | | CF | Controls | | | | |
| 2.5* | 7.56 | 12.6 ±4.5 | 10.8 ±2.9 | 83.0 | 100 | 88.2 | n.a. |
| 3 | 6.83 | 11.4 ±3.9 | 10.1 ±2.6 | 85.1 | 100 | 89.7 | 0.99 |
| 4 | 6.03 | 9.8 ±3.3 | 9.1 ±2.3 | 85.7 | 100 | 89.7 | 0.96 |
| 5 | 5.53 | 8.8 ±2.8 | 8.5 ±2.1 | 82.6 | 100 | 86.8 | 0.93 |
| 6 | 5.14 | 8.1 ±2.6 | 8.0 ±1.9 | 78.8 | 100 | 85.3 | 0.92 |
| 7 | 4.83 | 7.6 ±2.3 | 7.7 ±1.8 | 76.5 | 100 | 83.8 | 0.89 |
| 9 | 4.37 | 6.8 ±2.0 | 7.0 ±1.6 | 74.0 | 100 | 82.4 | 0.87 |
| 12 | 3.82 | 6.0 ±1.7 | 6.3 ±1.4 | 70.5 | 99.4 | 80.9 | 0.82 |
| 15 | 3.38 | 5.4 ±1.4 | 5.8 ±1.3 | 70.5 | 98.9 | 80.9 | 0.78 |
| 18 | 3.07 | 4.9 ±1.3 | 5.4 ±1.2 | 68.2 | 98.1 | 79.4 | 0.74 |

Lung clearance index (LCI) data from 68 children (44 with CF) with triplicate nitrogen multiple breath washout (N₂MBW) tests. Test duration (time in minutes) required to clear the lungs until respective target end-tidal N₂ concentration (cet) is given as mean (±SD). *2.5% N₂ cet of initial N₂ cet reflects the standard for LCI. Upper limits of normal LCI (ULN) is mean LCI + 1.96 SD from healthy controls. Sensitivity (Sens.), specificity (Spec.), and correctly classified (Correct) children are given in percent. R² gives the linear association of alternative LCI with standard LCI. n.a. = not applicable.

Table 2: Variability of standard and alternative lung clearance index.

| | Cystic Fibrosis | Healthy controls |
|-----------------------|------------------------|-------------------------|
| Standard LCI (TO) | 10.1 ±2.5 (6.5-15.9) | 6.4 ±0.6 (5.5-7.3) |
| Intra-test CV | 5.4% | 5.4% |
| LCI ₅ (TO) | 6.8 ±1.3 (4.8-9.4) | 4.9 ±0.3 (4.2-5.4) |
| Intra-test CV | 5.0% | 5.1% |

Standard lung clearance index (LCI) measured at 2.5% (1/40th) and at 5% (1/20th) of initial end-tidal nitrogen concentration differed significantly between 44 children with CF and 24 healthy controls ($p < 0.001$ for both). Data are given as mean ±SD (range) and are compared by unpaired t-test. LCI units are lung turnovers (TO). LCI intra-test coefficient of variation (CV) is derived from three nitrogen multiple-breath washout runs.

Table 3: Less than three - upper limit of normal, duration, and diagnostic accuracy of LCI.

| | ULN | Time (min) | | Sens. (%) | Spec. (%) | Correct (%) | R ² |
|-----------------------------|------|------------|-----------|-----------|-----------|-------------|----------------|
| | | CF | Controls | | | | |
| LCI_{single} | 7.66 | 2.2 ±1.0 | 1.7 ±0.6 | 84 | 97.5 | 89 | 0.97 |
| LCI_{double} | 7.58 | 7.4 ±2.9 | 6.2 ±1.8 | 87 | 100 | 92 | 0.99 |
| LCI_{std} | 7.56 | 12.6 ±4.5 | 10.8 ±2.9 | 84 | 100 | 89 | n.a. |

Lung clearance index (LCI) data from 68 children (44 with CF) based on one (LCI_{single}), two (LCI_{double}), and three (standard - LCI_{std}) nitrogen multiple breath washout (N₂MBW) tests. Test duration (time in minutes) required to clear the lungs until 1/40th of initial end-tidal nitrogen concentration is given as mean (±SD). Upper limits of normal LCI (ULN) is mean LCI + 1.96 SD from healthy controls. Sensitivity (Sens.), specificity (Spec.), and correctly classified (Correct) children are given in percent. R² gives the linear association of alternative LCI with standard LCI. n.a. = not applicable.

References

1. Stanojevic S, Wade A, Cole TJ, et al. Spirometry centile charts for young Caucasian children: the Asthma UK Collaborative Initiative. *Am J Respir Crit Care Med* 2009;**180**:547-552.
2. Robinson P, Latzin P, Verbanck S, et al. Guidelines for Inert Gas Washout Measurement using Multiple and Single Breath Tests [accepted]. *Eur Respir J* 2012.
3. Singer F, Houltz B, Latzin P, et al. A realistic validation study of a new nitrogen multiple-breath washout system. *PLoS ONE* 2012;**7**:e36083.
4. Singer F, Stern G, Thamrin C, et al. A new double-tracer gas single-breath washout to assess early cystic fibrosis lung disease. *Eur Respir J* 2012 May 17. doi: 10.1183/09031936.00044312. [Epub ahead of print]
5. Singer F, Stern G, Thamrin C, et al. Tidal volume single breath washout of two tracer gases - a practical and promising lung function test. *PLoS ONE* 2011;**6**:e17588.
6. Robinson P, Stocks J, Aurora P, et al. Abbreviated multi-breath washout for calculation of Lung Clearance Index. *Pediatr Pulmonol* 2012 Jul 25. doi: 10.1002/ppul.22618. [Epub ahead of print]
7. Gustafsson PM, Aurora P, Lindblad A. Evaluation of ventilation maldistribution as an early indicator of lung disease in children with cystic fibrosis. *Eur Respir J* 2003;**22**:972-979.
8. Fuchs SI, Ellemunter H, Eder J, et al. Feasibility and Variability of Measuring the Lung Clearance Index in a Multi-Center Setting. *Pediatr Pulmonol* 2012;**47**:649-657.
9. Aurora P, Gustafsson P, Bush A, et al. Multiple breath inert gas washout as a measure of ventilation distribution in children with cystic fibrosis. *Thorax* 2004;**59**:1068-1073.
10. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**1**:307-310.

11. Aurora P, Stanojevic S, Wade A, et al. Lung Clearance Index at 4 Years Predicts Subsequent Lung Function in Children with Cystic Fibrosis. *Am J Respir Crit Care Med* 2011;**183**:752-758.
12. Owens CM, Aurora P, Stanojevic S, et al. Lung Clearance Index and HRCT are complementary markers of lung abnormalities in young children with CF. *Thorax* 2011;**66**:481-488.
13. Hall GL, Logie KM, Parsons F, et al. Air Trapping on Chest CT Is Associated with Worse Ventilation Distribution in Infants with Cystic Fibrosis Diagnosed following Newborn Screening. *PLoS ONE* 2011;**6**:e23932.
14. Emmanuel G, Briscoe WA, Cournand A. A method for the determination of the volume of air in the lungs: measurements in chronic pulmonary emphysema. *J Clin Invest* 1961;**40**:329-337.
15. Kraemer R, Meister B. Fast real-time moment-ratio analysis of multibreath nitrogen washout in children. *J Appl Physiol* 1985;**59**:1137-1144.
16. Tsunoda S, Young AC, Martin CJ. Emptying pattern of lung compartments in normal man. *J Appl Physiol* 1972;**32**:644-649.
17. Green K, Buchvald FF, Marthin JK, et al. Ventilation inhomogeneity in children with primary ciliary dyskinesia. *Thorax* 2012;**67**:49-53.
18. Gustafsson PM, De Jong PA, Tiddens HA, et al. Multiple-breath inert gas washout and spirometry versus structural lung disease in cystic fibrosis. *Thorax* 2008;**63**:129-134.
19. Verbanck SA, Paiva M, Schuermans D, et al. Relationships between the lung clearance index and conductive and acinar ventilation heterogeneity. *J Appl Physiol* 2012;**112**:782-790.
20. Downie SR, Salome CM, Verbanck S, et al. Ventilation heterogeneity is a major determinant of airway hyperresponsiveness in asthma, independent of airway inflammation. *Thorax* 2007;**62**:684-689.