Bettina Frauchiger ORCID iD: 0000-0002-9519-9328

Marc-Alexander Oestreich ORCID iD: 0000-0001-9641-3691

Sophie Yammine ORCID iD: 0000-0001-7720-3445

Do clinimetric properties of LCI change after correction of signal processing?

Bettina Sarah Frauchiger, MD, PhD[1†] (https://orcid.org/0000-0002-9519-9328)

Marc-Alexander Oestreich, MD[1,2†] (https://orcid.org/0000-0001-9641-3691)

Florian Wyler[1] (https://orcid.org/0000-0002-1232-1392)

Nathalie Monney[1] (https://orcid.org/0000-0002-0627-6732)

Corin Willers, MD[1] (https://orcid.org/0000-0003-1140-9437)

Sophie Yammine, MD, PhD[1] (https://orcid.org/0000-0001-7720-3445)

Philipp Latzin, MD, PhD[1] (https://orcid.org/0000-0002-5239-1571)

[†] Authors contributed equally to this work as first author.

[1] Division of Paediatric Respiratory Medicine and Allergology, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland

[2] Graduate School for Health Sciences, University of Bern, Switzerland.

Corresponding author:

Prof. Dr. med. Philipp Latzin, PhD

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ppul.25865.

Inselspital, University Children's Hospital of Bern

Freiburgstrasse 15, 3010 Bern, Switzerland

E-mail: philipp.latzin@insel.ch

Phone: +41 31 632 9353

Keywords: Multiple-breath washout; Cystic fibrosis, pulmonary function testing

Running title: LCI properties hold after correction of sensor error

Word count: 3023/3500

Figures and tables: 5/6 (1 Table, 4 Figures)

References: 24/40

Online Supplement: This manuscript has supplemental material.

Abstract (226/250)

Background: The recently described sensor-crosstalk error in the multiple-breath washout (MBW) device Exhalyzer D (Eco Medics AG, Duernten, Switzerland) could highly influence clinimetric properties and the current interpretation of MBW results. This study reanalyzes MBW data from clinical routine in the corrected software version Spiroware® 3.3.1 and evaluates the effect on outcomes.

Methods: We included nitrogen-MBW data from healthy children and children with cystic fibrosis (CF) from previously published trials and ongoing cohort studies. We specifically compared lung clearance index (LCI) analyzed in Spiroware 3.2.1 and

3.3.1 with regards to i) feasibility, ii) repeatability and iii) validity as outcome parameters in children with CF.

Results: (i) All previously collected measurements could be reanalyzed and resulted in unchanged feasibility in Spiroware 3.3.1. (ii) Short- and midterm repeatability of LCI was similar in both software versions. (iii) Clinical validity of LCI remained similar in Spiroware 3.3.1, however, resulted in lower values. Discrimination between health and disease was comparable between both software versions. The increase in LCI over time was less pronounced with 0.16 LCI units/year (95% CI 0.08; 0.24) vs. 0.30 LCI units/year (95% CI 0.21; 0.38) in 3.2.1. Response to intervention in children receiving CFTR-modulator therapy resulted in a comparable improvement in LCI in both Spiroware versions.

Conclusion: Our study confirms that clinimetric properties of LCI remain unaffected after correction for the cross-sensitivity error in Spiroware software.

1. Introduction

We recently described and characterized a substantial sensor-crosstalk error in a commercially available and widely used multiple-breath washout (MBW) device (Exhalyzer D, Eco Medics AG, Duernten, Switzerland) and suggested a possible correction which is now available in an updated version of Spiroware analysis software (Spiroware 3.3.1, Eco Medics AG, Duernten, Switzerland) [1]. There is an obvious sense of concern in the MBW community as the potential impact of these findings on existing and ongoing studies could highly influence clinimetric properties of MBW results and their current clinical interpretation [2].

MBW has become an important diagnostic tool in cystic fibrosis (CF) for both, clinical follow-up of the patients and as an endpoint for clinical trials of new therapies [3-8]. The lung clearance index (LCI) is a feasible, repeatable, and sensitive marker of ventilation inhomogeneity that correlates well with structural lung disease and tracks disease progression in children with CF [5,6,9,10]. LCI is calculated as the ratio of the cumulative expired volume divided by the functional residual capacity (FRC), based on indirect calculation of nitrogen (N_2) from oxygen (O_2) and carbon dioxide (CO_2) signals. This results in inherent sensitivity to measurement errors in these signals, and the error described above previously resulted in a prolongation of the washout with falsely elevated LCI values [1].

To evaluate the potential clinical impact of these findings, this study aims to reanalyze published and collected MBW data from clinical routine by applying the corrected signal processing algorithm and evaluating the effect on outcomes. We assessed if previously described clinimetric properties of LCI hold for corrected results. We specifically examined corrected LCI with regards to i) feasibility, ii) repeatability, and iii) validity as outcome parameters in children with CF.

2. Methods

2.1 Study design and population

In this observational study combining several existing trials, we included N_2-MBW data from healthy children and children with CF from previously published trials and two ongoing cohort studies (Basel-Bern Infant Lung Development (BILD) cohort [11] and Swiss Cystic Fibrosis Infant Lung Development (SCILD) cohort [12]). The Ethics Committee of the Canton of Bern, Switzerland approved the study protocols (B2019-01072, PB_2017-02139, 2017-00279, 2017-00088) and parents gave written consent.

We included MBW data from 275 study participants that were reported in previous studies and 28 participants from an ongoing trial (Table 1). Using the corrected results of these measurements, we evaluated the effect of the correction on MBW outcomes and consequently if clinimetric properties of LCI were affected.

2.2 MBW measurements

N_2-MBW tests were performed using the Exhalyzer D device and Spiroware analysis software (v3.2.1, Eco Medics AG, Duernten, Switzerland) with weight-adjusted dead spaces and settings according to current consensus guidelines [13,14]. Testing was performed with patients sitting upright and breathing through a mouthpiece. Quality control was performed according to current guidelines and tests with at least two acceptable MBW trials were included in our analysis [14-16].

2.3 Correction of signal processing in MBW

In brief, the sensor-crosstalk error in the Exhalyzer D device underestimates O_2 and CO_2 gas concentrations and thus overestimates end-expiratory concentrations of N_2. Elevated N_2 concentrations cause a significant overestimation of the main MBW outcomes FRC and LCI as well as of the N_2 back-diffusion from the lungs[1]. An updated Spiroware analysis software (v3.3.1, Eco Medics AG, Duernten, Switzerland) will correct automatically for this error. Previously reported results, however, need to be reanalyzed. In this study, we reanalyzed raw data (A-Files) using LungSim 1.01, a custom Python script developed by our group replicating the signal processing and outcome calculation used in Spiroware analysis software (v3.2.1 and 3.3.1) [1].

2.4 Re-assessment of clinimetric properties and statistical analysis

We present non-parametric summaries for skewed data and parametric summaries for normally distributed characteristics. Statistical analyses were performed using Stata 16.1 (StataCorp 2019, College Station; TX). Figures were created using Stata 16.1 or Graph Pad Prism 8 (Prism G 2018, La Jolla; California). The following clinimetric properties of LCI were examined and compared between the Spiroware software versions 3.2.1 and 3.3.1:

2.4.1 feasibility (success rates of measurements)

Feasibility was defined by the number of study participants with acceptable MBW measurements, defined as at least two acceptable trials according to guidelines [13,14].

2.4.2 short- and mid-term repeatability (within 15 minutes and 24 hours)

As previously described, for short- and mid-term repeatability of LCI, MBW measurements were performed in triplicates 15 minutes and 24 hours apart respectively, with unchanged measurement conditions in school-aged children with CF [17,18]. We used mixed-effects linear regression models to calculate the variability in LCI between measurements allowing for correlation of repeated measurements within individuals. We used mean difference, intercept, and residual standard deviation to calculate the coefficient of variation (CV%), intraclass correlation coefficient (ICC), and coefficient of repeatability (CR). Based on a log-linear model, we calculated upper limit of normal (ULN) for relative differences (95% quantile of a normal distribution) indicating a threshold below which 95% of the relative differences are expected to fall.

2.4.3 validity

(a)discrimination between health and disease

For discrimination between health and disease, we used the ULN of LCI, calculated as mean LCI of healthy study participants + 1.96*Standard deviation (SD) from a previously described dataset [1].

b) correlation with functional MRI outcomes

Functional MRI data were acquired using matrix pencil decomposition MRI as previously described [18]. Main outcomes are fractional ventilation defect percentage (VDP) and relative perfusion defect percentage (QDP) expressed as the impaired fraction relative to whole lung volume. VDP and QDP were recalculated due to updated post-processing [19]. Correlations between VDP or QDP and LCI were assessed by Spearman's rank correlation [20].

c) longitudinal changes

Longitudinal changes in LCI were evaluated from a subset of the previously reported cohort including three monthly clinical routine MBW measurements in children with CF aged 3 - 18 years attending routine care between 2014-2018 in our center along with matching clinical information[5]. We used a mixed-effects linear regression model to assess the mean rate of change in LCI with age included as a linear term, a participant-specific random intercept and random slope to account for between-participant variability, different observation periods for each participant, and unequal numbers of study visits [21]. As previously described [5], we used a baseline model and a final model adjusted for predefined clinically most relevant covariates. Next, we

assessed all potentially influencing covariates on LCI course first in a univariate analysis and second in the fully adjusted model.

d) response to intervention

The response to intervention by modulator therapy was assessed by evaluating the change in LCI between baseline, i.e. before treatment start and after at least two weeks of treatment in three different treatment groups: In (i) either Orkambi® (Lumacaftor/Ivacaftor) or Symdeko® (Tezacaftor/Ivacaftor)-treated patients (double modulator therapy), in (ii) Trikafta® (Elexacaftor/Tezacaftor/Ivacaftor)–treated patients (triple modulator therapy), and (iii) patients who received first Orkambi® or Symdeko® followed by Trikafta® (combined modulator therapy). Response to intervention was defined as within-group changes from baseline to under treatment and analyzed by paired t-tests.

3. Results

1.1.1 3.1 Study population

Summaries for demographical characteristics and number of measurements included are presented in Table 1.

3.2.1 feasibility (success rates of measurements)

Due to the cross-sensitivity error, N_2 concentration was overestimated rather than underestimated, thus allowing the re-analysis of all previously collected measurements and resulting in unchanged feasibility. Following re-analysis, we found a reduction in CEV of 19.6% which results in a comparably shorter washout duration for the participant.

3.2.2 short- and mid-term repeatability

Short-term repeatability of LCI and FRC was assessed in sixteen children with CF and is summarized in Supplemental Table 1. Average LCI in children with CF was lower when analyzed in Spiroware 3.3.1 (LCI 9.4 (SD 1.8) vs. 10.8 (2.2)). In general, repeatability of MBW measurements was similar when analyzed in Spiroware 3.2.1 and 3.3.1. Variability between measurements expressed as CV% increased slightly in Spiroware 3.3.1 (LCI: 4.2 vs 3.7; FRC 3.7 vs 3.6). Also when expressing relative differences between measurements, the upper limit of normal (95% quantile) increased marginally (9.1 vs 8.4%).

Mid-term repeatability of LCI measurements was assessed in 12 healthy children and 23 children with CF, results are summarized in Supplemental Table 1. Mean (SD) LCI in healthy children was 6.8 (0.4) when assessed in Spiroware 3.2.1 and 6.1 (0.3) in Spiroware 3.3.1, in children with CF mean (SD) LCI was 11.6 (2.6) and 9.7 (2.2) in Spiroware 3.2.1 and 3.3.1, respectively. Repeatability indices were similar when assessed in Spiroware 3.2.1 and 3.3.1. While variability tended to be lower in healthy children when analyzed in Spiroware 3.3.1 (CV% 3.4 vs 4.4) this was the opposite in children with CF (CV% 9.6 (3.3.1) vs 8.1 (3.2.1). Similarly, when assessing relative changes between visits, the upper limit of normal (95% quantile) was lower in healthy children in Spiroware 3.3.1 compared to 3.2.1 but higher in children with CF in 3.3.1 (Supplemental Table 1). While we found within-subject between test standard deviation to be independent of the magnitude of LCI for both Spiroware settings (Supplemental Figure 1), within-subject within-test standard deviation was associated with the magnitude of LCI for both settings (Spiroware 3.2.1: R^2= 0.3 ($p<0.001$); 3.3.1: R^2= 0.2 ($p<0.001$).

3.2.3 validity

(a) discrimination between health and disease

ULN based on 75 healthy controls from a retrospective dataset [1] was lower in Spiroware 3.3.1 with LCI 7.1 compared to 8.1 in 3.2.1. Sensitivity was comparable (77.8% (3.2.1) and 76.2% (3.3.1)) and specificity identical (98.7%) between 3.2.1 and 3.3.1 with some scatter in the critical area (Figure 1, Supplemental Figure 2) based on natural variability. In Spiroware 3.3.1, two children with CF had abnormal LCI values compared to normal values in 3.2.1, while one healthy child was falsely categorized as abnormal based on slightly elevated LCI. In both Spiroware versions LCI values of all the other healthy controls were within normal range and 15 children with CF had values below the ULN.

b) correlation with functional MRI outcomes

We reanalyzed the correlation of functional MRI and LCI in 14 children with CF. The extent of QDP ranged between 15% and 35% (Figure 2) The correlation between QDP and LCI remained consistently strong in both Spiroware versions (3.2.1: r_S=0.66, p=0.03; 3.3.1: r_S=0.69 p=0.02). The increase in LCI per increase in ventilation defect was comparable for both Spiroware versions (Spiroware 3.2.1: 1.28 LCI units/percent perfusion defect (95% CI 0.67; 1.89); Spiroware 3.3.1: 1.32 LCI units/percent perfusion defect (95% CI 0.55; 2.10); Supplemental Figure 3).

c) longitudinal tracking

To assess the differences in the longitudinal course of LCI between Spiroware 3.2.1 and 3.3.1, we reanalyzed 796 measurements from 72 children with CF (Table 1). Without adjustment for risk factors, LCI increase was comparable in both settings,

however less pronounced in Spiroware 3.3.1 (0.16 LCI units/year (95% CI 0.08; 0.24) vs. 0.30 LCI units/year (95% CI 0.21; 0.38) in Spiroware 3.2.1). The pattern of increase in LCI was similar in both settings, remaining stable during preschool years and school-age and then starting to increase in adolescence (Figure 3 and Supplemental Table 2). Similar to our previous findings, Aspergillus and P. aeruginosa colonization, severe exacerbations, and experiencing ABPA during the study period remained individually associated with a steeper increase in LCI also in Spiroware 3.3.1, even though with a smaller magnitude (Supplemental Table 3). The effect on covariates associated with acute changes in LCI (acute exacerbations, CF-related diabetes, BMI z-score) remained similar for both Spiroware settings (Supplemental Table 4). With adjustment for previously defined risk factors (sex, BMI, PsA- and Aspergillus-colonization, CF-related diabetes, acute and severe exacerbations) the pattern of increase in LCI was similar for both Spiroware settings, again, less pronounced for 3.3.1 (0.08 LCI units/year (95% CI 0.01; 0.14) (Spiroware 3.3.1) vs. 0.19 LCI units/year (95% CI 0.12; 0.27) in Spiroware 3.2.1) (Supplemental Figure 4).

d) Response to intervention To characterize differences in the response to intervention with double or triple modulator therapy between Spiroware 3.2.1 and 3.3.1, we reanalyzed 212 visits from 28 patients (Table 1) and compared mean LCI values at baseline, under double, under triple, and combined therapy. There was a statistically significant improvement (reduction) in LCI in all three treatment groups when compared to baseline in both Spiroware algorithms (Figure 4). In Spiroware 3.2.1, within-group mean (95% CI) absolute change from baseline was -1.7 LCI units (-2.8 to -0.5, p=0.012) under double therapy, -1.7 LCI units (-2.5 to -0.9; $p \leq 0.001$) under triple therapy, and -2.5 LCI units (-4.1 to -1.0; p=0.007) under combined

modulator therapy (Figure 4 and supplemental Table 5). In Spiroware 3.3.1, LCI was substantially lower but the change from baseline remained statistically significant in all groups (within-group mean (95% CI) absolute change from baseline was -1.5 LCI units (-2.5 to -0.4, p=0.013) under double therapy, -1.3 LCI units (-1.9 to -0.7; p≤0.001) under triple therapy, and -1.7 LCI units (-2.9 to -0.5; p=0.014) under combined modulator therapy). Overall, within-group mean values differed substantially between Spiroware 3.2.1 and 3.3.1 (mean (95% CI) difference between software versions over all groups was -1.5 (-2.0 to -1.1) LCI units).

4. Discussion

4.1 Summary

In this study, we can confirm that clinimetric properties of LCI are still valid after correction for the recently described measurement error in the widely used MBW Spiroware software [1]. As expected, the correction led to lower LCI values and thus lower thresholds, implicating that we need to re-define existing thresholds and accordingly adjust clinical interpretation of LCI and its changes.

4.2 Comparison to literature

4.2.1 Clinical impact of measurement error

The initial concerns that the findings of Wyler et al. [1] about the measurement error in the widely used MBW software Spiroware 3.2.1 could tremendously change utility of LCI in clinical routine, can be rejected by our findings. We could show, that clinimetric properties of LCI still hold, even though with a slightly smaller magnitude based on overall lower LCI values. This is due to the measurement principle of the test, the indirect N_2-calculation was found to be overestimated by the cross-sensitivity

error in the O_2 and CO_2- signals, leading to a prolongation of the washout with falsely high LCI values. Our results suggest that clinical interpretation of LCI change is not affected by the correction of this error. As such we can support recently published results from clinical trials and extent the interpretation into the clinical application [2].

4.2.2 Impact on feasibility (success rates of measurements)

Since the correction of the gas measurement error results in shorter washouts, we were able to reanalyze all our trials with the corrected software version [22]. Thus, also in clinical practice, overall N_2-MBW will become shorter with the corrected software which might increase the feasibility of the test, especially in the younger and possibly less cooperative age range but also in patients with more advanced disease and longer washout times [13,22].

4.2.3 Impact on Repeatability

Overall, we found short-and midterm repeatability [17,18] to be similar after cross-sensitivity correction, and thus limits for intervisit changes and variability to be unaffected. Within-subject between test standard deviation remained independent of the magnitude of LCI, which preserves the property of LCI to guide individual clinical decisions.

4.2.4 Validity

The discriminatory ability of LCI holds after correction with a linear trend to lower LCI values for healthy but also children with CF. Thereby, the higher the LCI value, the higher the downward correction, which needs closer reevaluation of patients with highly elevated LCI. Potentially, the overestimation of back-diffusion of N_2 might have overrated LCI values of patients with more severe disease and longer washouts. However, correlation with functional defects detected by MRI scans [17,18,20] remained

very similar before and after measurement correction, reassuring that the pathophysiological understanding of LCI is still valid. Correspondingly, the pattern of increase in LCI with age remained similar with a smaller magnitude after measurement correction (Figure 3, Supplemental Figure 4). A very similar effect was seen by Robinson et al. reanalyzing change in LCI in preschool CF children over 12 months [2]. In line with Robinson et al., treatment effects of CFTR modulators remained statistically significant whereas the magnitude of the change was smaller.

4.3 Strengths and weaknesses of this study

We performed a thorough reanalysis of clinical and research MBW data using the corrected algorithm incorporated in Spiroware 3.3.1, with data originally collected or reloaded in Spiroware 3.2.1 [23]. We performed rigorous quality control with only perfect data being used [15] which minimizes bias due to sampling variation. Further, we were able to use a wide range of different datasets obtained in a clinical routine setting or from clinical studies to assess the impact of measurement correction on various clinimetric properties of LCI. By nature, the main limitation lies in the retrospective application of the correction. Besides, we had two different datasets from healthy controls available, one consisting of twelve subjects providing repeated LCI measurements 24h apart to assess repeatability and the other consisting of 75 subjects providing one LCI measurement to calculate normative values. The inherently wider spread of the second dataset along with the different statistical approach in the first dataset accounting for repeated measures led to a small difference in the ULN for LCI (first dataset 6.7 vs. 7.1 in the second dataset, Spiroware 3.3.1).

4.4 Clinical relevance and outlook

Although there is a significant change with the corrected algorithm, we show that the clinimetric properties of LCI still hold. This means that clinical utility and our understanding of LCI and its changes remain valid. This also suggests that previously published papers do not need to be withdrawn. However, this implies also that for absolute correct estimation of effect sizes, previously published datasets need to be reanalyzed to enable comparison with results from Spiroware 3.3.1. Besides, we will need to closely evaluate changes within individuals applying the new software, to adjust individual therapeutic goals but also to capture individual disease progression at this new level. The option of retrospective correction has the great advantage that all previously collected data in Spiroware can be reanalyzed and corrected.

5. Conclusion

Our study confirms that clinimetric properties of LCI remain unaffected after correction for the recently detected measurement error in the widely used Spiroware software. However, the correction results in lower LCI values, with the concurrent need to redefine existing thresholds and adjust clinical understanding of LCI and its changes on this new level.

6. Acknowledgements

We thank all our patients and study participants as well as their families for allowing their MBW data to be used for research. In addition, we thank the lab technicians and study nurses for their contribution by performing MBW measurements. Furthermore, we thank the physicians from our department, especially Insa Korten and Elisabeth Kieninger.

7. Funding

This project was funded by the Swiss National Science Foundation, Grant Nos. 182719 (to P. Latzin) and 179905 (to S. Yammine).

8. Author contributions

Bettina Sarah Frauchiger: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Writing - Original Draft, Review & Editing, Visualization.

Marc-Alexander Oestreich: Conceptualization, Data Curation, Methodology, Formal analysis, Investigation, Resources, Writing - Original Draft, Review & Editing, Visualization.

Florian Wyler: Software.

Nathalie Monney: Resources.

Corin Willers: Resources.

Sophie Yammine: Writing - Original Draft, Review & Editing.

Philipp Latzin: Conceptualization, Supervision, Project administration, Funding acquisition.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

1. Wyler F, Oestreich MA, Frauchiger BS, Ramsey KA, Latzin P. 2021. Correction of sensor crosstalk error in exhalyzer d multiple-breath washout device significantly impacts outcomes in children with cystic fibrosis. Journal of applied physiology (Bethesda, Md: 1985). 131(3):1148-1156.

2. Robinson PD, Jensen R, Seeto RA, Stanojevic S, Saunders C, Short C, Davies JC, Ratjen F. 2021. Impact of cross-sensitivity error correction on representative nitrogen-based multiple breath washout data from clinical trials. Journal of Cystic Fibrosis.

3. Amin R, Stanojevic S, Kane M, Webster H, Ratjen F. 2016. A randomized controlled trial to evaluate the lung clearance index as an outcome measure for early phase studies in patients with cystic fibrosis. Respiratory medicine. 112:59-64.

4. Davies G, Stanojevic S, Raywood E, Duncan JA, Stocks J, Lum S, Bush A, Viviani L, Wade A, Calder A et al. 2020. An observational study of the lung clearance index throughout childhood in cystic fibrosis: Early years matter. The European respiratory journal.2000006.

5. Frauchiger BS, Binggeli S, Yammine S, Spycher B, Kruger L, Ramsey KA, Latzin P. 2021. Longitudinal course of clinical lung clearance index in children with cystic fibrosis. The European respiratory journal. 58(1):2002686.

6. Perrem L, Stanojevic S, Shaw M, Jensen R, McDonald N, Isaac SM, Davis M, Clem C, Guido J, Jara S et al. 2020. Lung clearance index to track acute

respiratory events in school-age children with cystic fibrosis. Am J Respir Crit Care Med.

7. Rayment JH, Stanojevic S, Davis SD, Retsch-Bogart G, Ratjen F. 2018. Lung clearance index to monitor treatment response in pulmonary exacerbations in preschool children with cystic fibrosis. Thorax. 73(5):451-458.

8. Voldby C, Green K, Kongstad T, Ring AM, Sandvik RM, Skov M, Buchvald F, Pressler T, Nielsen KG. 2020. Lung clearance index-triggered intervention in children with cystic fibrosis - a randomised pilot study. Journal of cystic fibrosis: official journal of the European Cystic Fibrosis Society. 19(6):934-941.

9. Oude Engberink E, Stanojevic S, Ratjen F. 2016. Clinimetric properties of the lung clearance index in adults and children with cystic fibrosis. Chest. 150(6):1412-1413.

10. Ramsey KA, Rosenow T, Turkovic L, Skoric B, Banton G, Adams A-M, Simpson SJ, Murray C, Ranganathan SC, Stick SM et al. 2016. Lung clearance index and structural lung disease on computed tomography in early cystic fibrosis. Am J Respir Crit Care Med. 193(1):60–67.

11. Fuchs O, Latzin P, Kuehni CE, Frey U. 2012. Cohort profile: The bern infant lung development cohort. Int J Epidemiol. 41(2):366–376.

12. Korten I, Kieninger E, Yammine S, Regamey N, Nyilas S, Ramsey K, Casaulta C, Latzin P, For TSSG. 2018. The swiss cystic fibrosis infant lung development (scild) cohort. Swiss Med Wkly. 148:w14618.

13. Robinson PD, Latzin P, Ramsey KA, Stanojevic S, Aurora P, Davis SD, Gappa M, Hall GL, Horsley A, Jensen R et al. 2018. Preschool multiple-breath washout testing. An official american thoracic society technical statement. Am J Respir Crit Care Med. 197(5):e1-e19.

14. Robinson PD, Latzin P, Verbanck S, Hall GL, Horsley A, Gappa M, Thamrin C, Arets HGM, Aurora P, Fuchs SI et al. 2013. Consensus statement for inert gas washout measurement using multiple- and single- breath tests. Eur Respir J. 41(3):507–522.

15. Frauchiger BS, Carlens J, Herger A, Moeller A, Latzin P, Ramsey KA. 2021. Multiple breath washout quality control in the clinical setting. Pediatr Pulmonol. 56(1):105-112.

16. Jensen R, Stanojevic S, Klingel M, Pizarro ME, Hall GL, Ramsey K, Foong R, Saunders C, Robinson PD, Webster H et al. 2016. A systematic approach to multiple breath nitrogen washout test quality. PLoS One. 11(6):e0157523.

17. Roethlisberger K, Nyilas S, Riedel T, Wolfensberger J, Singer F, Latzin P. 2018. Short-term effects of elastic chest wall restriction on pulmonary function in children with cystic fibrosis. Respiration. 96(6):535-542.

18. Nyilas S, Bauman G, Pusterla O, Ramsey K, Singer F, Stranzinger E, Yammine S, Casaulta C, Bieri O, Latzin P. 2019. Ventilation and perfusion assessed by functional mri in children with cf: Reproducibility in comparison to lung function. Journal of Cystic Fibrosis. 18(4):543-550.

19. Willers C, Bauman G, Andermatt S, Santini F, Sandkühler R, Ramsey KA, Cattin PC, Bieri O, Pusterla O, Latzin P. 2021. The impact of segmentation on whole-lung functional mri quantification: Repeatability and reproducibility from multiple human observers and an artificial neural network. Magn Reson Med. 85(2):1079-1092.

20. Nyilas S, Bauman G, Sommer G, Stranzinger E, Pusterla O, Frey U, Korten I, Singer F, Casaulta C, Bieri O et al. 2017. Novel magnetic resonance technique for functional imaging of cystic fibrosis lung disease. The European respiratory journal. 50(6):1701464.

21. Grégoire TG, Schabenberger O, Barrett JP. 1995. Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. Canadian Journal of Forest Research. 25(1):137-156.

22. Yammine S, Summermatter S, Singer F, Lauener R, Latzin P. 2016. Feasibility of nitrogen multiple-breath washout in inexperienced children younger than 7 years. Pediatr Pulmonol. 51(11):1183-1190.

23. Bayfield KJ, Shaar A, Robinson PD. 2021. Further considerations on normative data for multiple breath washout outcomes. Eur Respir J. 57(4):2004536.

Figures

Figure 1. Discrimination between health and disease holds after correction for cross-sensitivity error. Scatter plot of LCI values for healthy controls (hollow circles) and participants with CF (solid black) obtained in Spiroware 3.2.1 and 3.3.1. The dashed lines represent upper limits of normality (95% quantile). Abbreviations: HC: healthy controls, CF: cystic fibrosis, LCI: lung clearance index, TO: turnover, ULN: upper limit of normal (95% quantile).

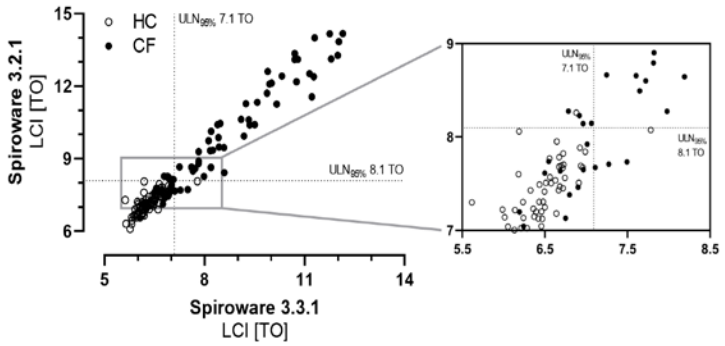

Figure 2. Association of perfusion defect from functional MRI and LCI remains similar after correction for cross-sensitivity error. Shown are the associations between the perfusion defect [%] from functional MRI and the LCI [TO] in Spiroware 3.2.1 (A) and Spiroware 3.3.1 (B) for n=14 CF patients. Abbreviations: CF: cystic fibrosis, LCI: lung clearance index, TO: turnover.

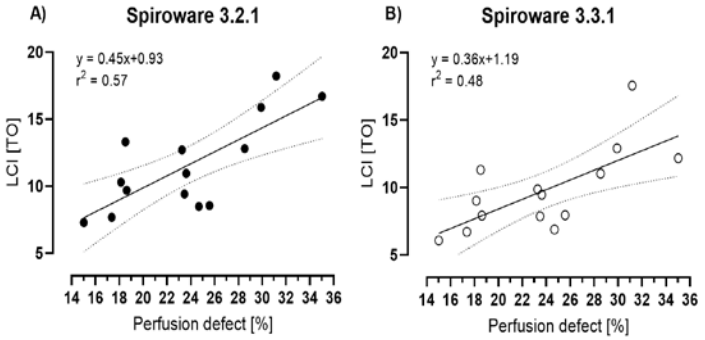

Figure 3. LCI increases over time without adjustments in Spiroware 3.2.1 and 3.3.1. In Spiroware 3.2.1 (A)**,** the increase in LCI over age was 0.30 LCI units/year (95% CI 0.21; 0.38). Increase in preschool-age was -0.26 LCI units/year (95%CI -0.95; 0.44), in school-age 0.22 LCI units/year (95%CI 0.08; 0.35), in adolescence 0.40 LCI units/year (95%CI 0.26; 0.54) (p-value for interaction 0.02), with a steeper increase in females during adolescence. In Spiroware 3.3.1 (B)**,** the increase in LCI

over age was 0.16 LCI units/year (95% CI 0.08; 0.24). Increase in preschool-age was -0.27 LCI units/year (95%CI -0.87,0.33), in school-age 0.11 LCI units/year (95%CI -0.01; 0.22), in adolescence 0.30 LCI units/year (95%CI 0.18; 0.43) (p-value for interaction 0.02), also with a steeper increase in females during adolescence. On the y-axis, LCI raw values are given. The solid line represents mean LCI values across all participants with available data at a given age. Shaded areas represent point-wise upper and lower 95% confidence intervals, the dotted line refers to an upper limit of normal for LCI of 8 TO. Abbreviations: LCI: lung clearance index, TO: turnover, CI: Confidence interval.

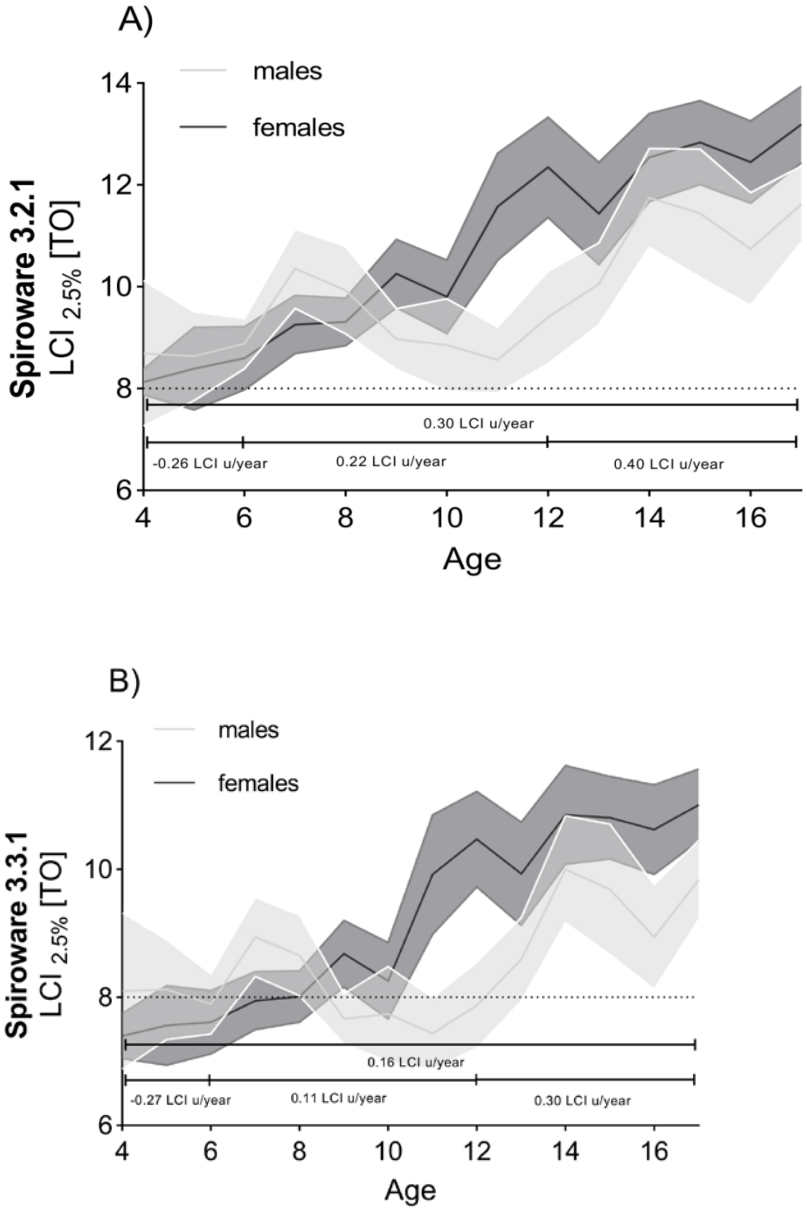

Figure 4. Responsiveness to intervention with modulator therapy holds after correction for cross-sensitivity error. Data are presented as mean (95% CI) and before-after plots. Shown are mean (95%) LCI values from 212 visits of 28 patients at baseline, under double (Orkambi® (n=8)), under triple (Trikafta® (n=13)), and under combined therapy (first double followed by triple) modulator therapy (Orkambi® (n=2) or Symdeko® (n=5) followed by Trikafta®). Statistical analysis: paired t-test (baseline vs. endpoint). Abbreviations: LCI: lung clearance index, TO: turnover, 95% CI: 95% confidence interval.

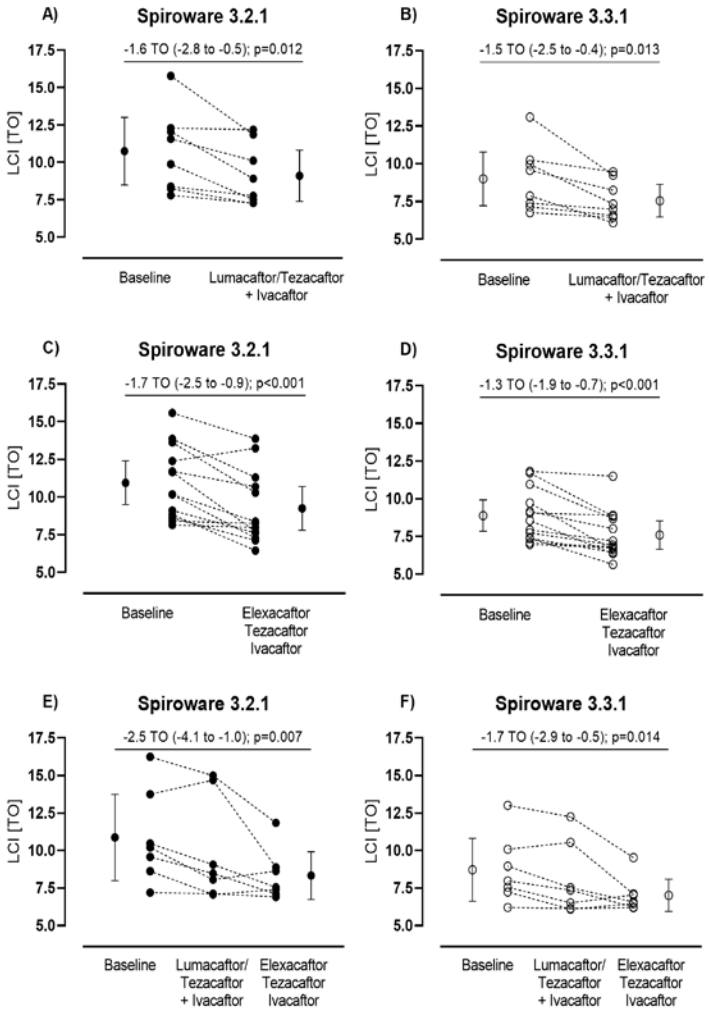

Table 1. Demographic characteristics of study population

	Repeatability		Validity			
	short-term (15 min.)	mid-term (24 hours)	Discrimination between health and disease	Longitudinal evolution of LCI*	Correlation with functional MRI	**Response to intervention**
Patients (n)	16	35	138	72	14	28
Healthy	-	12	75	-	-	-
Cystic fibrosis	16	23	63	72	14	28
Age (years)	12.4 (6.6; 17.2)	12.5 (5.6; 18.1)	10.0 (5.6; 18.1)	9.6 (4; 17.3)	14.0 (6.1; 18.9)	12.4 (5.7; 16.4)
Weight (kg)	40.0 (13.1)	41.7 (15.2)	41.7 (15.2)	31.1 (12.9)	45.4 (13.9)	39.1 (10.8)
Length (cm)	145.1 (14.4)	147.0 (18.8)	147.0 (18.8)	132.7 (20.2)	154.9 (16.2)	147.6 (14.3)

Table 1. Demographic characteristics of the study populations for included datasets. Data are presented as mean (SD). Age is shown as mean (min to max). *Values for the longitudinal evolution of LCI are baseline values. Abbreviations: MRI: magnetic resonance imaging, LCI: lung clearance index.