

1 **Correction of measurement error in a commercial multiple-breath washout**
2 **device**

3 Florian Wyler*¹ (<https://orcid.org/0000-0002-1232-1392>)

4 Marc-Alexander Oestreich*^{1,2} (<https://orcid.org/0000-0001-9641-3691>)

5 Bettina S. Frauchiger^{1,3} (<https://orcid.org/0000-0002-9519-9328>)

6 Kathryn Ramsey^{#1} (<https://orcid.org/0000-0003-4574-6917>)

7 Philipp Latzin^{#1} (<https://orcid.org/0000-0002-5239-1571>)

8 * Authors contributed equally to this work as first author

9 # Authors contributed equally to this work as last author

10

11 ¹ Pediatric Respiratory Medicine, Inselspital, University Children's Hospital of Bern, University
12 of Bern, Switzerland.

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

14 ³ Graduate School of Cellular and Biomedical Sciences, University of Bern, Switzerland

15

16 **Abstract**

17 Rationale:

18 Nitrogen multiple-breath washout (N₂MBW) is an established technique to assess functional
19 residual capacity (FRC) and ventilation inhomogeneity in the lung. Accurate measurement of
20 gas concentrations is essential for the appropriate calculation of clinical outcomes.

21 Objectives:

22 We investigated the accuracy of oxygen and carbon dioxide measurements used for the
23 indirect calculation of nitrogen concentration in a commercial MBW device (Exhalyzer D, Eco
24 Medics AG, Duernten, Switzerland) and its impact on FRC and lung clearance index (LCI).

25 Methods:

26 High precision calibration gas mixtures and mass spectrometry were used to evaluate sensor

27 output. We assessed the impact of corrected signal processing on FRC and LCI in a dataset
28 of healthy children and children with cystic fibrosis using custom analysis software.

29 Results:

30 We found inadequate correction for the cross sensitivity of the oxygen and carbon dioxide
31 sensors in the Exhalyzer D device. This results in an overestimation of expired nitrogen
32 concentration, and consequently FRC and LCI outcomes. Breath-by-breath correction of this
33 error reduced mean (SD) FRC by 8.9 (2.2)% and LCI by 11.9 (4.0)%. It also resulted in almost
34 complete disappearance of the tissue nitrogen signal at the end of measurements.

35 Conclusions:

36 Inadequate correction for cross sensitivity between the oxygen and carbon dioxide gas sensors
37 of the Exhalyzer D device leads to an overestimation of FRC and LCI. Correction of this error
38 is possible and could be applied by re-analysing the measurements breath-by-breath in an
39 updated software version.

40

41 **Corresponding Author**

42 Prof. Philipp Latzin, MD, PhD

43 Inselspital, University Children's Hospital of Bern

44 Freiburgstrasse 8, 3010 Bern, Switzerland

45 E-mail: philipp.latzin@insel.ch

46 Phone: +41 31 632 9353

47

48 **Author contributions**

49 F.W. and M.O. designed performed and analysed experiments. M.O., B.F and K.R created and
50 organized the databases for retrospective analysis. F.W. wrote the manuscript, with major
51 contributions from M.O, K.R. and P.L. All authors checked the final version of the manuscript.

52

53 **Grants, Gifts, Equipment, Drugs**

54 Eco Medics AG (Duernten, Switzerland) provided a research version of their commercial
55 software Spiroware 3.2.1 including insight on signal processing algorithms and helped with
56 the acquisition of mass spectrometry measurements. This project was funded by the Swiss
57 National Science Foundation, Grant Nr. 182719 (P. Latzin) and 168173 (K. Ramsey)

58 **Introduction**

59 Nitrogen multiple-breath washout (N₂MBW) is an established technique to assess ventilation
60 inhomogeneity and lung volumes during relaxed tidal breathing¹⁻³. The test starts at the end of
61 a relaxed expiration, with the lung at functional residual capacity (FRC). As subjects inhale
62 pure oxygen (100% O₂), resident nitrogen (N₂) is gradually washed out. The main outcome,
63 the lung clearance index (LCI), represents the expired volume required to reach 2.5% of the
64 starting N₂ concentration expressed in multiples of the FRC.

65 The N₂MBW technique is increasingly being used as a functional measure capable of
66 monitoring early disease progression more sensitive than spirometry and to capture treatment
67 effects in patients with conditions such as cystic fibrosis (CF) and primary ciliary dyskinesia
68 (PCD)⁴⁻⁸. The endorsement of the North American Cystic Fibrosis Foundation and the
69 European Cystic Fibrosis Society Clinical Trials Network consolidated the LCI as an endpoint
70 in clinical trials in children and adults with CF^{3,9-12}. As recommended by different societies,
71 most data on the efficacy of disease modifying therapies in CF^{13,14} were obtained with the
72 Exhalyzer D / Spiroware setup (Eco Medics AG, Duernten, Switzerland) and this is the device
73 of choice in multiple ongoing studies with the LCI as a primary endpoint (www.clinicaltrials.gov:
74 NCT04138589, NCT03320382, NCT04026360, NCT02657837). Consequently, the correct
75 measurement of N₂MBW is of critical importance, notably in studies that serve to guide drug
76 approval. In this device N₂ concentration is not measured directly, but relies on indirect
77 determination of nitrogen, through accurate measurement of oxygen (O₂) and carbon dioxide
78 (CO₂)¹⁵.

79 In this study we investigated the accuracy of indirect measurement of N₂ during MBW tests
80 using the Exhalyzer D device by i) assessing the sensor accuracy in the Exhalyzer D's O₂ and
81 CO₂ sensors, ii) establishing a correction for any observed sensor error and iii) assessing the
82 effect size of sensor error on clinical outcomes (FRC, LCI) as well as on tissue nitrogen.

83 **Methods**

84 **Study design**

85 This was a combined prospective, experimental and retrospective, observational study to
86 assess sensor accuracy as well as the impact of sensor inaccuracy using existing N₂MBW
87 data. We included experimental data from gas mixture measurements and mass spectrometry
88 (Eco Medics AG, Duernten, Switzerland). We characterized the impact of the developed
89 correction functions on MBW outcomes with data from healthy children¹⁶ and children
90 diagnosed with CF (Swiss Cystic Fibrosis Infant Lung Development (SCILD) cohort)^{6,17}. The
91 Ethics Committee of the Canton of Bern, Switzerland approved the study protocol (PB_2017-
92 02139).

93 **i) Sensor accuracy**

94 To measure sensor accuracy over the wide range of concentrations encountered in a N₂MBW
95 measurement, we tested the Exhalyzer D sensors with gas mixtures containing known
96 reference concentrations, and compared concentrations of CO₂ and O₂ measured by the
97 Exhalyzer D to known reference concentrations (supplemental Table 1). Specifications from
98 manufacturers of technical gas mixtures (Carbagas AG, Muri bei Bern, Switzerland) and mass
99 spectrometry measurements (Eco Medics AG, Duernten, Switzerland) served as reference.

100 **Equipment**

101 Applying Dalton's law of partial pressures, the Exhalyzer D (Eco Medics AG, Duernten,
102 Switzerland) computes an indirect nitrogen concentration that is based on the measurement
103 of oxygen, carbon dioxide and an assumed fixed ratio of nitrogen and argon (supplemental
104 Equation 1)¹⁸. While dry atmospheric air contains nitrogen (N₂), oxygen (O₂), argon, carbon
105 dioxide (CO₂) and traces of neon, helium, methane and krypton, it is assumed that all gases
106 except oxygen and carbon dioxide (which are affected by tidal breathing) remain in proportion
107 to each other.

108 **Sensor characteristics**

109 The O₂ sensor used in the Exhalyzer D setup (X3004 OEM sensor, Oxigraf Inc., Sunnyvale,
110 CA, USA) is located within the main structure of the Exhalyzer D. Gas from the patients
111 breathing stream is sampled at a rate of 200 mL/min via a Nafion tube¹⁸. The CO₂ sensor
112 (Capnostat 5, Respironics Inc., Wallingford, CT, USA) is located within the breathing stream.
113 The Exhalyzer D measures both O₂ and CO₂ using laser absorption spectroscopy, a technique
114 where the absorption of light is measured at specific frequencies characteristic to each gas¹⁹.
115 The absorption spectrum of both gases is affected by a variety of factors, including pressure,
116 temperature, and the presence of other gases (manufacturer's manual). The presence of O₂
117 in the gas mixture affects the absorption spectrum of CO₂ and vice versa.

118 **Software for gas sensor measurements**

119 The manufacturer provided a research version of the Exhalyzer D software (Spiroware 3.3.0
120 Research, Eco Medics AG, Duernten, Switzerland) which enabled technical gas
121 measurements with direct recording of raw, uncorrected sensor outputs. This software version
122 is, for the purpose of MBW tests, identical with the commercial release (Spiroware 3.2.1) on
123 the level of signal processing and analysis.

124 **Technical gases**

125 Sensor accuracy was tested over the complete range of concentrations present in N₂MBW
126 measurements. Twelve technical gas mixtures (Carbagas AG, Muri bei Bern, Switzerland),
127 each containing different combinations of CO₂ (2.5%, 5%, 7.5%) and O₂ (30%, 60%, 90%) as
128 well as three mixtures containing only CO₂ and O₂, were used (supplemental Table 1).
129 Technical gas mixtures were manufactured at 2% mixture precision and 1% measurement
130 precision. Additionally, a series of mass spectrometry measurements was carried out, where
131 N₂ was kept at 2% to mimic the MBW end of test condition, while CO₂ varied from 0-6% with
132 the rest of the mixture being O₂.

133 **Measurement conditions for technical gases and mass spectrometry**

134 Sensor accuracy of the Exhalyzer D was assessed under two conditions: i) at body
135 temperature ($37.5 \pm 1.5^\circ \text{C}$), humidity saturated (MR850 Heated Humidifier, Fisher & Paykel,
136 Auckland, New Zealand), with technical gas mixture specifications as reference, and ii) at 35°
137 C in a dry, climate-controlled chamber with mass spectrometry (AMIS 2000, Innovision ApS,
138 Odense, Denmark and Red-y Smart Controller, Voegtlin Instruments GmbH, Muttenz,
139 Switzerland) as reference. Gas concentrations were measured in triplicates, on two Exhalyzer
140 D devices (calibrated according the manufacturers manual) for technical gas measurements
141 and one mass spectrometry device. Additional calibration measurements were carried out
142 before and after each set of measurements.

143 **Signal processing for gas sensors**

144 The CO_2 signal measured was subsequently corrected with the relevant signal processing of
145 a standard Exhalyzer D measurement to achieve final sensor readings that were equivalent to
146 current standard processing during a MBW measurement: For condition i) a
147 temperature/humidity-dependant "ATP correction" factor was applied to the CO_2 -signal to
148 correct for humidity and transform the measured fraction into atmospheric temperature and
149 pressure (ATP) conditions (Eq. 2 in OLS). For both condition i) and ii), the CO_2 signal was
150 corrected with an O_2 -dependant cross sensitivity correction factor (Eq. 3 in OLS).

151 **ii) Correction function fitting**

152 We combined the data of the technical gas mixture and mass spectrometry measurements to
153 construct a correction function for the O_2 and CO_2 sensors. Sensor error was defined as the
154 absolute difference between measured concentrations of O_2 and CO_2 by the Exhalyzer D and
155 reference values provided by mass spectrometry and the technical gas mixture manufacturer.
156 We then fitted a 2nd-degree two-parameter polynomial through the error for each sensor, as a
157 function of measured O_2 and CO_2 (Eq. 4 in OLS): Fitting was performed using MATLAB 2017b
158 (Mathworks, Natick, Massachusetts, USA). This characterization of the measurement error as
159 a function of measured gas concentrations could then be directly used as a correction function

160 for the analysis of MBW measurements, by adding the fitted error for each combination of O₂
161 and CO₂ to the respective measured concentrations.

162 **iii) Effect size of sensor correction**

163 **Retrospective analysis**

164 We characterized the impact of measurement error on MBW outcomes using a custom Python
165 script developed by our group, designed to replicate standard Spiroware 3.2.1 signal
166 processing and outcome calculation of MBW data. The impact of sensor error was assessed
167 by comparing the outcomes of standard analysis (Exhalyzer D and Spiroware 3.2.1, Eco
168 Medics AG, Duernten, Switzerland) with the outcomes of a corrected algorithm modified to
169 include the O₂-CO₂ cross sensitivity correction functions outlined above.

170 For this, we re-analysed 357 MBW measurements of 85 healthy controls (HC, mean age:
171 11.1y, SD: 3.8y, range: 6-17.8y) and 62 subjects diagnosed with cystic fibrosis (CF, mean age:
172 9.5y, SD: 3.5y, range: 4-17.3y) from previous studies^{6,16} (OLS Table 2 for details). We included
173 study visits with at least two acceptable, quality controlled N₂-MBW measurements, in
174 accordance with recent consensus guidelines²⁰.

175 **Statistical testing**

176 For statistical testing of significance we used STATA 16 (StataCorp LLC, College Station,
177 Texas, USA). Unpaired t-tests were used for comparisons between healthy controls and
178 children with cystic fibrosis. Paired t-tests were performed for differences between standard
179 and corrected processing of MBW files.

180 **Results**

181 **i) Sensor accuracy**

182 **O₂ sensor accuracy**

183 We found that the O₂ sensor of the Exhalyzer D device has a non-linear O₂- and CO₂-
184 dependant measurement error (Figure 1). The degree of error in the O₂ sensor is higher with
185 increasing concentrations of CO₂ between the range of 0 – 7.5%. At conditions that reflect the
186 end of the current standard MBW end of test (N₂ around 2%, expired CO₂ around 5%, rest O₂),
187 the O₂ sensor error leads to an underestimation of absolute expiratory O₂-concentrations of
188 0.8%. This in turn leads to an overestimation of calculated concentrations of N₂, and makes up
189 the majority (86%) of a total sensor error which in the example above leads to measured
190 concentrations of N₂ of 2% instead of the corrected 1.03% (Table 1).

191 **CO₂ sensor accuracy**

192 Additionally we found a residual error in the CO₂ sensor, currently not corrected for. The CO₂
193 sensor reading is already corrected by a factor that depends on the concentration of O₂ (see
194 methods). However, the CO₂ sensor seems to display a different cross-sensitivity
195 measurement error than the current signal processing takes into account (Figure 2). This
196 residual error appears to be mainly dependant on CO₂ concentration (Figure 2A) and only
197 mildly dependant on O₂ concentration (Figure 2B). Overall, the impact of this residual error in
198 CO₂ is lower than the error of the O₂ sensor. It also leads to an overestimation of N₂, making
199 up 14% of the total sensor error in the standard end of test condition (Table 1).

200 The combined error of both sensors is given in Figure 3. As illustrated, the resulting error in N₂
201 is higher in very low N₂ concentrations and at higher CO₂ concentrations. The gas
202 concentrations that exist at a typical end of N₂MBW measurements (corrected N₂ around 2%
203 and CO₂ around 5%) unfortunately result in a relatively large absolute N₂-error of 0.94% and
204 even larger relative N₂-error of 47.2% (Table 1). The lowered N₂ concentrations throughout the
205 entire MBW measurement (Figure 4A) lead to an end of test that is systematically reached
206 earlier (Figure 4B).

207 **iii) Effect size of sensor correction**

208 **Sensor correction impact on MBW outcomes**

209 We re-analysed 357 MBW measurements from healthy controls (HC) and children with cystic
210 fibrosis (CF) using the sensor correction function outlined above in a custom software. Applying
211 the O₂ and CO₂ sensor correction function had a significant impact on all MBW outcomes
212 (Table 1). Following the sensor correction, mean (SD) FRC and LCI decreased by 8.9 (2.2) %
213 and 11.9 (4.0) %, respectively. The cumulative expired volume (breathing required by the
214 patient, CEV) decreased by 19.6 (5.0) %. The reduced CEV is explained by lower end-
215 expiratory concentrations of N₂, which lead to an earlier end of test. Decreased FRC is
216 explained by slightly lower concentrations of N₂ throughout the measurement. The decrease
217 in CEV is more pronounced than for FRC, and with LCI being the ratio of those two outcomes
218 ($LCI = CEV/FRC$), this leads to an overall decrease in LCI.

219 Outcome differences due to the sensor correction can vary greatly for individual measurements
220 (Table 2), and without re-analysis on a signal processing level it is difficult to predict how much
221 outcomes of one measurement will change due to the sensor correction. However, outcomes
222 on the level of a larger number of measurements before and after the correction correlate
223 strongly. Linear fits of corrected outcomes vs standard outcomes have R² values of 0.99 for
224 FRC, and 0.96 for LCI (Figure 5).

225 The significance of differences in LCI and FRC [L] observed between healthy controls and
226 children with CF in standard processing were preserved after sensor correction (Table 3). On
227 the other hand, the difference in FRC [mL/kg] between groups only became significant after
228 correction.

229 The change in outcomes following correction is dependent on the magnitude of the outcomes
230 themselves, for both FRC and LCI (Figure 6).

231 **Sensor correction impact on tissue nitrogen**

232 Due to the fact that N₂ is currently overestimated in the presence of CO₂, there is a non-zero
233 N₂ concentration calculated by the Exhalyzer D, even without any N₂ present. This minimum
234 measurable N₂ concentration in current standard processing can be estimated by summing up

235 the O₂ and CO₂ sensor correction functions (Figure 7B). In conditions reflecting the end of
236 expirations in conditions where N₂ is absent after correction (CO₂ around 5%, rest O₂), the
237 measured concentration of N₂ by the Exhalyzer D in standard processing is 0.94% (Table 1,
238 Figure 7B, intersection of green line with x-axis). For this given concentration of CO₂, the
239 measured N₂ concentration using an Exhalyzer D could theoretically never reach lower end-
240 expiratory values than this. If the Exhalyzer D is therefore used to quantify the impact of tissue
241 nitrogen in long MBW measurements, a significant part of what appears to be a diffusion
242 equilibrium of tissue nitrogen disappears after the sensor correction is applied (Figure 7A). The
243 higher the end-expiratory concentrations of CO₂, the greater the measured (artificial)
244 concentration of N₂ when real N₂ is at 0%. (Figure 7B).

245 **Discussion**

246 **Summary**

247 We report a significant measurement error in the Eco Medics Exhalyzer D N₂MBW device. At
248 high concentrations of O₂, and natural end-expiratory concentrations of CO₂, the Exhalyzer D
249 sensors underestimate O₂ and CO₂ gas concentrations and therefore overestimate end-
250 expiratory concentrations of N₂. This strongly affects the end of test criterion, and causes
251 significant error in the assessment of FRC and LCI. It also creates a significant artificial
252 overestimation of measured tissue nitrogen at the end of MBW tests.

253 **i) Sensor accuracy**

254 We are the first group to identify this sensor cross sensitivity error. Previous validation studies
255 as well as the internal testing of Ecomedics did not make special mention of the end of test,
256 end-expiratory conditions examined here¹⁵. It is worth noting that this error remains relatively
257 constrained for the readings of the individual sensors themselves (ca.1% relative error of a
258 sensor reading) even in the most extreme case (low N₂, high CO₂). However, due to the design
259 of the Exhalyzer D MBW device, and due to the fact that high O₂ and high CO₂ concentrations
260 occur at the end of the MBW measurement, this sensor error translates into a relative N₂ error
261 in the range of 47% at the corrected end of test condition. This measurement error greatly
262 exceeds the recommendations for manufacturers outlined in the ATS/ERS consensus
263 statement of measuring within relative 5% of tracer gas concentration²⁰.

264 **ii) Correction function**

265 The sensor error we observed in this study appears systematic and reproducible across
266 different Exhalyzer D systems, with multiple different ground truth references confirming our
267 findings. The correction function required to correct for the sensor error is simple and can be
268 implemented in future signal processing of Spiroware, as well as applied retrospectively to
269 existing data.

270 **iii) Effect size of sensor correction**

271 **Sensor correction impact on MBW outcomes**

272 The current erroneous calculated N₂ reading in the end of test end-expiratory condition leads
273 to substantially lowered MBW outcomes. This result provides a new perspective on previously
274 described differences between N₂ and SF6 MBW measurements²¹. It also offers a potential
275 explanation for the differences observed between N₂MBW outcomes measured using the
276 Exhalyzer D and devices by other manufacturers such as ndd Medizintechnik AG (Zürich,
277 Switzerland)²². In both studies, it was observed that primary MBW outcomes from the
278 Exhalyzer D were higher than in the other systems, an observation that may be explained by
279 the systematic overestimation of N₂ by the Exhalyzer D reported in this study. The direction of
280 the change after correction suggests that differences between devices will be smaller. In order
281 to validate this in detail, original data need to be reloaded using the sensor correction described
282 here. Fortunately, the N₂ error has been an overestimation rather than an underestimation, as
283 measurements can now be re-analysed without the worry that the trials might not have
284 recorded data long enough to reach the end of test in the corrected measurement.

285 A major concern that arises with the publication of this study is that it calls into question
286 previously published results obtained using the Exhalyzer D. It is to be expected that effect
287 sizes, confidence intervals and significance values of MBW outcomes in such studies will
288 change. However, while the impact of the sensor error has effects which are difficult to predict
289 on the level of individual measurements, the impact on MBW outcomes on a large enough
290 number of files appears more systematic. It is therefore reasonable to be optimistic that
291 publications which observed significant differences in MBW outcomes between two groups or
292 as a treatment effect may continue to see a difference after the sensor correction re-analysis
293 – even though the changes observed here suggest that effect sizes will be smaller. In our
294 opinion, re-analysis of MBW measurements should become a priority for those studies where
295 e.g. drug approval was or is based on affected N₂MBW data.

296 **Sensor correction impact on tissue nitrogen**

297 One previously-described feature of N₂ MBW tests is the fact that towards the end of a N₂MBW
298 test the concentration of N₂ in the lungs drops so low that a noticeable amount of N₂ diffuses
299 from the body into the lungs^{21,23}. The results of this study would suggest that the impact of
300 tissue N₂ is significantly lower than previously measured with Exhalyzer D devices. The error
301 measured in the end-expiratory, high-O₂ condition seems to suggest that the Exhalyzer D could
302 not have measured any end-expiratory (CO₂ around 5%) concentrations of N₂ lower than
303 0.94%, which would significantly perturb estimates of tissue nitrogen. The sensor correction
304 functions introduced in this paper would therefore eliminate a substantial part of the observed
305 tissue nitrogen in measurements performed with the Exhalyzer D (approximately 1% N₂ at
306 equilibrium).

307 **Strengths and limitations**

308 The main strength of this study is the fact that we were able, through detailed understanding
309 of the underlying signal processing of the Exhalyzer D, to characterize the precise impact of
310 an observed error in gas sensors on the clinical outcomes LCI and FRC. The findings from the
311 technical gases were confirmed by measurements using a Mass Spectrometer, both giving
312 similar results. We developed a correction function and were thus able to estimate the
313 measurement error precisely and correct for it.

314 The main limitation of this study is the fact that we only had a finite number of gas samples
315 with finite precisions to test the sensors with. We chose a selection of gas concentrations from
316 our range of interest which would exhibit cross-sensitivity effects, but could ultimately not cover
317 the entire range of concentration combinations in MBW measurements using technical gases.
318 However, the phase of the measurement where sensor accuracy is the most relevant for
319 accurate MBW outcomes is the end of test, where thanks to mass spectrometry measurements
320 we can describe the sensor error with high certainty.

321 Due to the minor contribution of argon (1.2% of N₂) and the negligible contribution of neon,
322 helium, methane, krypton etc.) these gases were omitted in the gas mixtures. In this study, the
323 portion of gases other than oxygen and carbon dioxide is therefore referred to as nitrogen.

324 This does not change anything about the relative errors observed, although in examples closer
325 to the reality of an MBW test, 1.2% of the fraction attributed here to N₂ would be considered
326 Argon.

327 **Outlook**

328 In the process of conducting the research for this paper, we contacted the manufacturer for
329 information regarding their sensor configurations and questions regarding sensor settings and
330 signal processing. They are working on a solution to both re-analyse past measurements and
331 to improve the signal processing for future measurements in an upcoming version of
332 Spiroware.

333 **Conclusion**

334 An error in the cross sensitivity correction between the oxygen and carbon dioxide gas sensors
335 of the Exhalyzer D device leads to an overestimation of FRC and LCI. Correction of this error
336 is possible but needs to be applied breath-by-breath by re-analysing the measurements in an
337 updated software version.

338 **References**

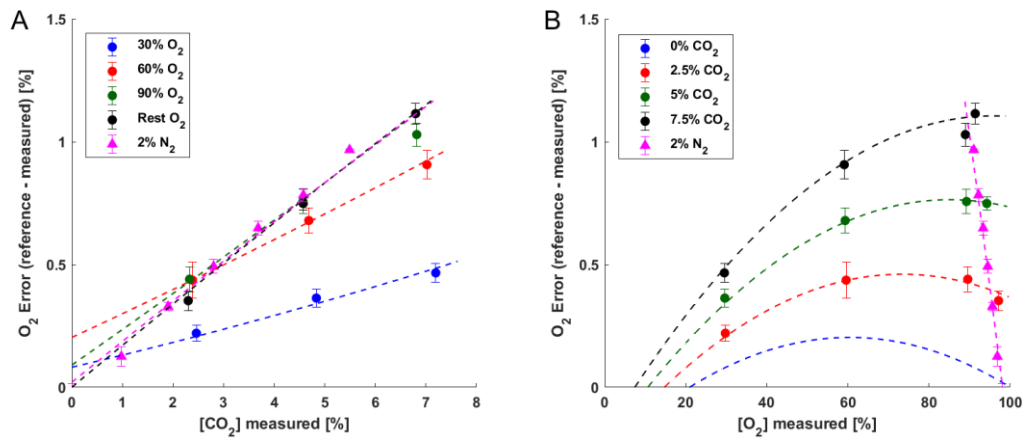
- 339 1. Usemann J, Yammine S, Singer F, Latzin P. Inert gas washout: background and
340 application in various lung diseases. *Swiss Med Wkly*. 2017;147:w14483.
341 doi:10.4414/smw.2017.14483.
- 342 2. Robinson PD, Goldman MD, Gustafsson PM. Inert gas washout: theoretical background
343 and clinical utility in respiratory disease. *Respiration*. 2009;78(3):339–355.
344 doi:10.1159/000225373.
- 345 3. Kent L, Reix P, Innes JA, Zielen S, Le Bourgeois M, Braggion C, Lever S, Arets HGM,
346 Brownlee K, Bradley JM, et al. Lung clearance index: evidence for use in clinical trials in
347 cystic fibrosis. *J Cyst Fibros*. 2014;13(2):123–138. doi:10.1016/j.jcf.2013.09.005.
- 348 4. Perrem L, Stanojevic S, Shaw M, Jensen R, McDonald N, Isaac SM, Davis M, Clem C,
349 Guido J, Jara S, et al. Lung Clearance Index to Track Acute Respiratory Events in
350 School-age Children with Cystic Fibrosis. *Am J Respir Crit Care Med*. 2020.
351 doi:10.1164/rccm.202006-2433OC.
- 352 5. Stanojevic S, Davis SD, Retsch-Bogart G, Webster H, Davis M, Johnson RC, Jensen R,
353 Pizarro ME, Kane M, Clem CC, et al. Progression of Lung Disease in Preschool Patients
354 with Cystic Fibrosis. *Am J Respir Crit Care Med*. 2017;195(9):1216–1225.
355 doi:10.1164/rccm.201610-2158OC.
- 356 6. Frauchiger BS, Binggeli S, Yammine S, Spycher B, Krüger L, Ramsey KA, Latzin P.
357 Longitudinal Course of Clinical Lung Clearance Index in Children with Cystic Fibrosis.
358 *Eur Respir J*. 2020. doi:10.1183/13993003.02686-2020.
- 359 7. Nyilas S, Bauman G, Pusterla O, Sommer G, Singer F, Stranzinger E, Heyer C, Ramsey
360 K, Schlegtendal A, Benzrath S, et al. Structural and Functional Lung Impairment in
361 Primary Ciliary Dyskinesia. Assessment with Magnetic Resonance Imaging and Multiple-
362 breath Washout in Comparison to Spirometry. *Ann Am Thorac Soc*. 2018;15(12):1434–
363 1442. doi:10.1513/AnnalsATS.201712-967OC.
- 364 8. Kinghorn B, McNamara S, Genatossio A, Sullivan E, Siegel M, Bauer I, Clem C, Johnson
365 RC, Davis M, Griffiths A, et al. Comparison of Multiple-breath Washout and Spirometry in

- 366 Children with Primary Ciliary Dyskinesia and Cystic Fibrosis and Healthy Controls. *Ann*
367 *Am Thorac Soc.* 2020. doi:10.1513/AnnalsATS.201905-375OC.
- 368 9. Saunders C, Jensen R, Robinson PD, Stanojevic S, Klingel M, Short C, Davies JC,
369 Ratjen F. Integrating the multiple-breath washout test into international multicentre trials.
370 *J Cyst Fibros.* 2020;19(4):602–607. doi:10.1016/j.jcf.2019.11.006.
- 371 10. Subbarao P, Milla C, Aurora P, Davies JC, Davis SD, Hall GL, Heltshe S, Latzin P,
372 Lindblad A, Pittman JE, et al. Multiple-Breath Washout as a Lung Function Test in Cystic
373 Fibrosis. A Cystic Fibrosis Foundation Workshop Report. *Ann Am Thorac Soc.*
374 2015;12(6):932–939. doi:10.1513/AnnalsATS.201501-021FR.
- 375 11. Milla CE, Ratjen F, Marigowda G, Liu F, Waltz D, Rosenfeld M. Lumacaftor/Ivacaftor in
376 Patients Aged 6-11 Years with Cystic Fibrosis and Homozygous for F508del-CFTR. *Am*
377 *J Respir Crit Care Med.* 2017;195(7):912–920. doi:10.1164/rccm.201608-1754OC.
- 378 12. Davies J, Sheridan H, Bell N, Cunningham S, Davis SD, Elborn JS, Milla CE, Starner TD,
379 Weiner DJ, Lee P-S, et al. Assessment of clinical response to ivacaftor with lung
380 clearance index in cystic fibrosis patients with a G551D-CFTR mutation and preserved
381 spirometry: a randomised controlled trial. *Lancet Respir Med.* 2013;1(8):630–638.
382 doi:10.1016/S2213-2600(13)70182-6.
- 383 13. Davies JC, Sermet-Gaudelus I, Naehrlich L, Harris RS, Campbell D, Ahluwalia N, Short
384 C, Haseltine E, Panorchan P, Saunders C, et al. A phase 3, double-blind, parallel-group
385 study to evaluate the efficacy and safety of tezacaftor in combination with ivacaftor in
386 participants 6 through 11 years of age with cystic fibrosis homozygous for F508del or
387 heterozygous for the F508del-CFTR mutation and a residual function mutation. *J Cyst*
388 *Fibros.* 2020. doi:10.1016/j.jcf.2020.07.023.
- 389 14. Ratjen F, Hug C, Marigowda G, Tian S, Huang X, Stanojevic S, Milla CE, Robinson PD,
390 Waltz D, Davies JC, et al. Efficacy and safety of lumacaftor and ivacaftor in patients aged
391 6–11 years with cystic fibrosis homozygous for F508del-CFTR: a randomised, placebo-
392 controlled phase 3 trial. *The Lancet Respiratory Medicine.* 2017;5(7):557–567.
393 doi:10.1016/S2213-2600(17)30215-1.

- 394 15. Singer F, Houltz B, Latzin P, Robinson P, Gustafsson P. A realistic validation study of a
395 new nitrogen multiple-breath washout system. *PLoS ONE*. 2012;7(4):e36083.
396 doi:10.1371/journal.pone.0036083.
- 397 16. Anagnostopoulou P, Latzin P, Jensen R, Stahl M, Harper A, Yammine S, Usemann J,
398 Foong RE, Spycher B, Hall GL, et al. Normative data for multiple-breath washout
399 outcomes in school-aged Caucasian children. *Eur Respir J*. 2020;55(4).
400 doi:10.1183/13993003.01302-2019.
- 401 17. Korten I, Kieninger E, Yammine S, Regamey N, Nyilas S, Ramsey K, Casaulta C, Latzin
402 P, For TSSG. The Swiss Cystic Fibrosis Infant Lung Development (SCILD) cohort. *Swiss
403 Med Wkly*. 2018;148:w14618. doi:10.4414/smw.2018.14618.
- 404 18. Gustafsson PM, Robinson PD, Lindblad A, Oberli D. Novel methodology to perform
405 sulfur hexafluoride (SF₆)-based multiple-breath wash-in and washout in infants using
406 current commercially available equipment. *J Appl Physiol (1985)*. 2016;121(5):1087–
407 1097. doi:10.1152/jappphysiol.00115.2016.
- 408 19. Cummings B, Hamilton ML, Ciaffoni L, Pragnell TR, Peverall R, Ritchie GAD, Hancock
409 G, Robbins PA. Laser-based absorption spectroscopy as a technique for rapid in-line
410 analysis of respired gas concentrations of O₂ and CO₂. *J Appl Physiol (1985)*.
411 2011;111(1):303–307. doi:10.1152/jappphysiol.00119.2011.
- 412 20. Robinson PD, Latzin P, Verbanck S, Hall GL, Horsley A, Gappa M, Thamrin C, Arets
413 HGM, Aurora P, Fuchs SI, et al. Consensus statement for inert gas washout
414 measurement using multiple- and single- breath tests. *Eur Respir J*. 2013;41(3):507–522.
415 doi:10.1183/09031936.00069712.
- 416 21. Jensen R, Stanojevic S, Gibney K, Salazar JG, Gustafsson P, Subbarao P, Ratjen F.
417 Multiple-breath nitrogen washout: a feasible alternative to mass spectrometry. *PLoS
418 ONE*. 2013;8(2):e56868. doi:10.1371/journal.pone.0056868.
- 419 22. Poncin W, Singer F, Aubriot A-S, Lebecque P. Agreement between multiple-breath
420 nitrogen washout systems in children and adults. *J Cyst Fibros*. 2017;16(2):258–266.
421 doi:10.1016/j.jcf.2016.11.004.

- 422 23. Nielsen N, Nielsen JG, Horsley AR. Evaluation of the impact of alveolar nitrogen
423 excretion on indices derived from multiple-breath nitrogen washout. PLoS ONE.
424 2013;8(9):e73335. doi:10.1371/journal.pone.0073335.

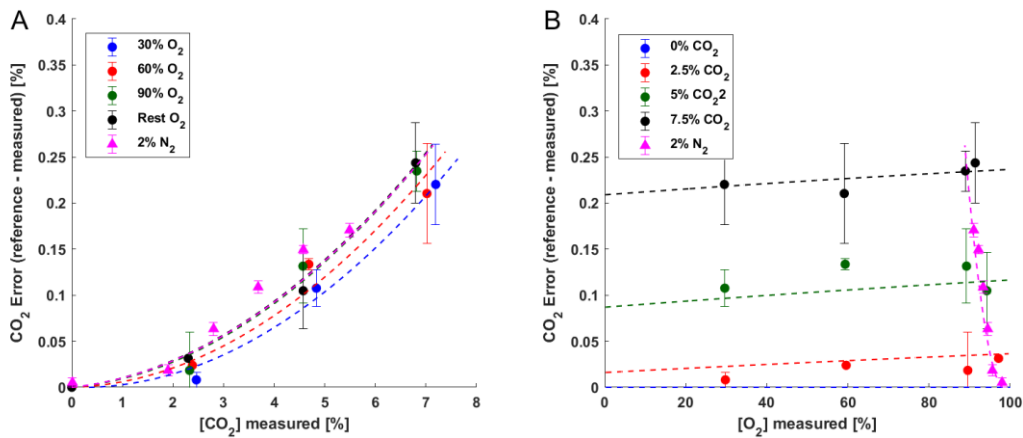
425 **Figure 1**



426

427 *Figure 1: Observed absolute error between known reference and measured O₂ concentrations. (dots: mean of error*
428 *for one gas mixture, error bars: +/- SD of error). Curves represent a two parameter quadratic polynomial fitted*
429 *through the error values (see OLS for details), represented here as curves for given CO₂ or O₂ concentrations. Dots*
430 *represent measurements performed with 12 technical gas mixtures as reference, triangles represent mass*
431 *spectrometry reference measurements of 6 mixtures at the end of test condition (2% N₂). **A)** Absolute O₂ error as*
432 *a function of CO₂ concentration, **B)** Absolute O₂ error as a function of O₂ concentration. Dashed curves all represent*
433 *the same fit $f(O_2, CO_2)$.*

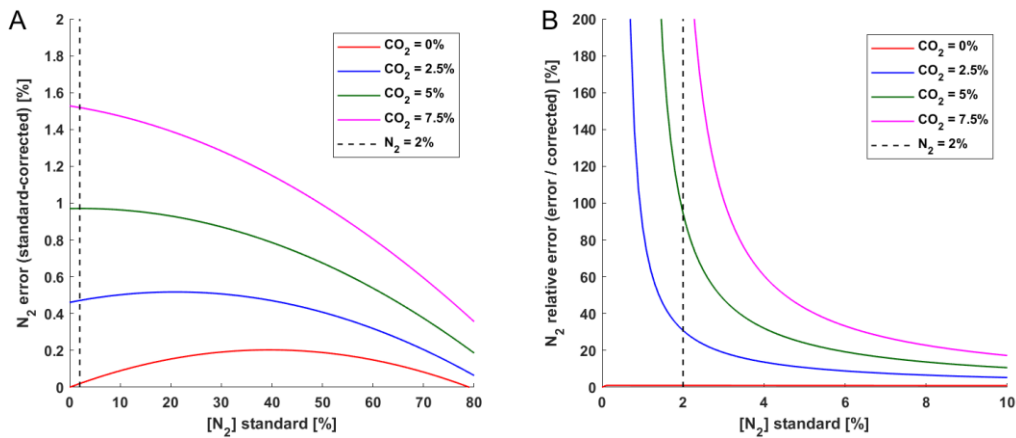
434 **Figure 2**



435

436 *Figure 2: Observed absolute error between known and measured CO₂ concentrations. (dots: mean of error of one*
437 *gas mixture, error bars: +/- SD of error). Curves represent a two parameter quadratic polynomial fitted through the*
438 *error values (see OLS for details), represented here as curves for given CO₂ or O₂ concentrations. Dots represent*
439 *measurements performed with 12 technical gas mixtures as reference, triangles represent mass spectrometry*
440 *reference measurements on 6 mixtures at the end of test condition (2% N₂). **A**) Absolute CO₂ error as a function of*
441 *CO₂ concentration, **B**) Absolute CO₂ error as a function of O₂ concentration. Dashed curves all represent the same*
442 *fit $f(O_2, CO_2)$.*

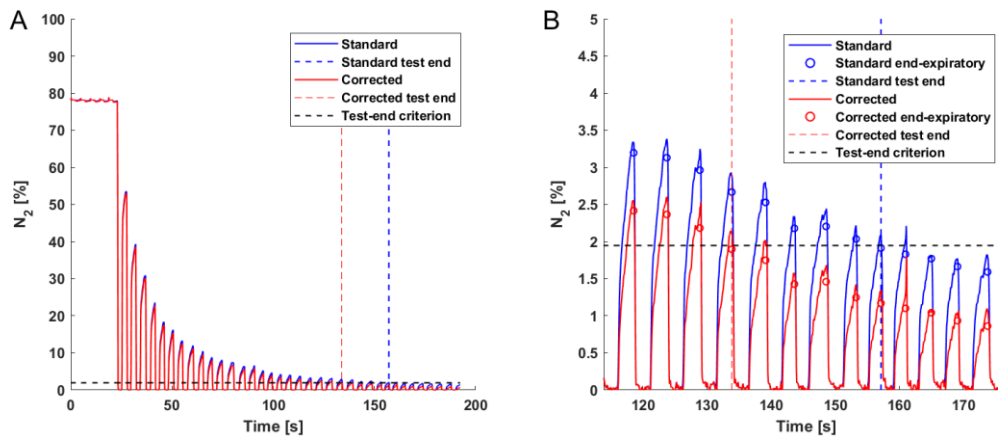
443 **Figure 3**



444

445 *Figure 3: N₂ error as a function of measured N₂. Illustration of the impact of the O₂ and CO₂ correction functions*
446 *on final measured N₂. **A)** Summed up correction functions of Figure 1 and Figure 2, and therefore absolute error in*
447 *N₂, represented as curve for selected concentrations of CO₂. **B)** Relative error in measurement of N₂ (standard-*
448 *corrected)/corrected) as a function of N₂.*

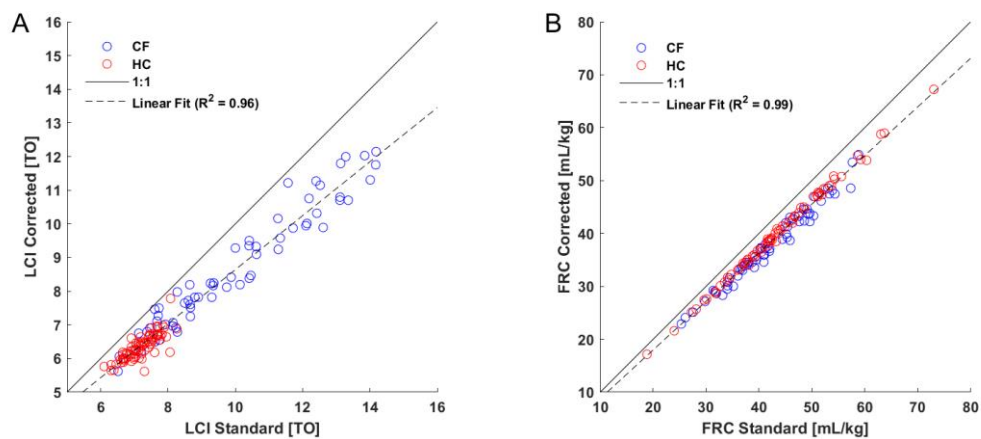
449 **Figure 4**



450

451 *Figure 4: Illustration of the effect of the sensor correction on the N_2 signal and consequently on the end of test of*
452 *an example MBW measurement. Traced in blue is the signal output of the standard signal processing, corrected*
453 *signal is shown in red. Vertical dashed lines represent the end of test for the current standard and corrected*
454 *measurement respectively. The black line corresponds to $1/40^{\text{th}}$ of the initial N_2 concentration (end of test criterium).*
455 **A)** *Time course of N_2 throughout a standard MBW measurement. B)* *Zoom into the critical period of end of test*
456 *determination. In this example the test ends 5 breaths earlier in the corrected measurement compared to standard.*

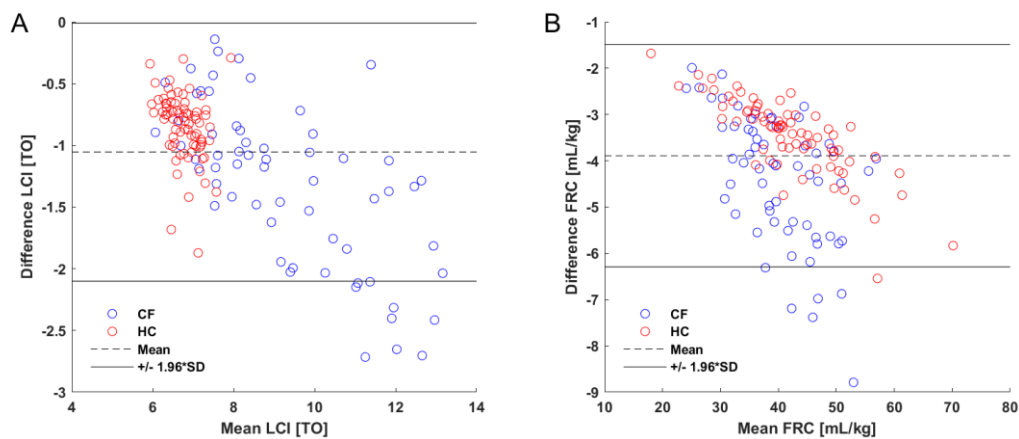
457 **Figure 5**



458

459 *Figure 5: MBW outcomes (A) LCI (B) FRC) after sensor correction (corrected) vs current standard (standard), in*
460 *healthy controls (HC) and patients with cystic fibrosis (CF).*

461 **Figure 6**

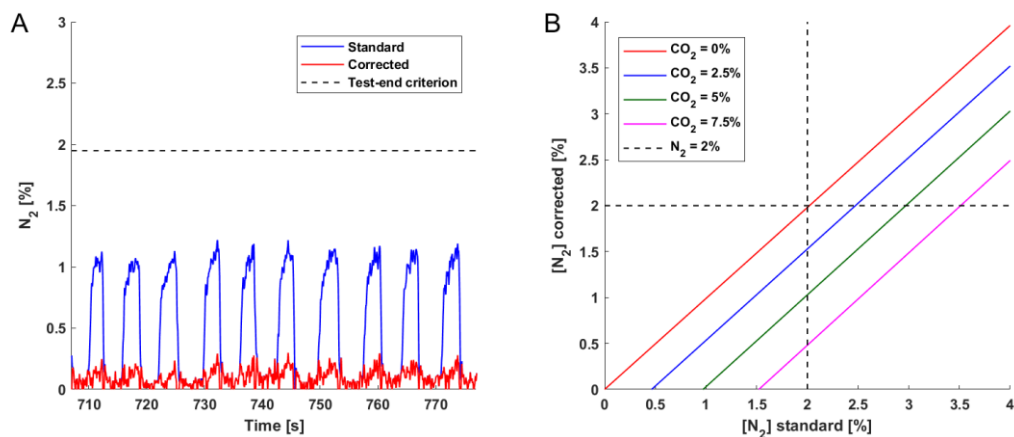


462

463 *Figure 6: Bland-Altman plot of the absolute difference (corrected – standard) in MBW outcomes (A) LCI (B) FRC)*

464 *due to sensor correction, plotted as a function of the outcomes themselves (mean of corrected and standard).*

465 **Figure 7**



466

467 *Figure 7: Illustration of the effect of the sensor correction function on nitrogen measurement in the late phase of*
468 *MBW tests. A) Example of the equilibrium N_2 reached in a very long continued MBW measurement using the*
469 *Exhalyzer D, displaying a greatly decreased N_2 -back-diffusion equilibrium (tissue nitrogen). B) Corrected N_2*
470 *plotted against standard N_2 in conditions around the end of test condition (2% N_2).*

471 **Table 1**

| Condition | Standard | | | Corrected | | | N ₂ Error | | Contribution | |
|------------------------|-------------------|--------------------|-------------------|-------------------|--------------------|-------------------|----------------------|------|--------------------|-------------------|
| | [N ₂] | [CO ₂] | [O ₂] | [N ₂] | [CO ₂] | [O ₂] | abs | rel | [CO ₂] | [O ₂] |
| Standard test-end [%] | 2 | 5 | 93 | 1.03 | 5.14 | 93.8 | 0.97 | 94.4 | 14.2 | 85.8 |
| Corrected test-end [%] | 2.94 | 4.87 | 92.2 | 2 | 5 | 93 | 0.94 | 47.2 | 13.9 | 86.1 |
| No nitrogen [%] | 0.94 | 4.87 | 94.2 | 0 | 5 | 95 | 0.94 | | | |

472

473 *Table 1: Specific examples of sensor error impact on measurement of N₂ in three conditions of interest. Standard*
474 *concentrations denote concentrations measured in standard Spiroware 3.2.1 processing. Corrected concentrations*
475 *correspond to concentrations after sensor correction is applied. N₂ error summarizes the absolute difference*
476 *between N₂ in standard vs. corrected, as well as the relative error (standard-corrected)/corrected. The relative*
477 *contribution of each sensor to the total error is listed under "Contribution". For simplicity, the presence of Argon is*
478 *omitted here.*

479 **Table 2**

| | n | Standard | Corrected | Difference | | | | |
|-------------|-----|----------|-----------|------------|----------|--------|---------------|-------------------|
| | | | | mean | mean [%] | SD [%] | 95% CI [%] | p-value |
| FRC [mL/kg] | 147 | 43.9 | 40.0 | 3.9 | -8.9 | 2.2 | -9.3 - -8.6 | <0.0001 |
| CEV [L] | 147 | 14.9 | 11.9 | 3.0 | -19.6 | 5.0 | -20.4 - -18.8 | <0.0001 |
| LCI [TO] | 147 | 8.3 | 7.3 | 1.0 | -11.9 | 4.0 | -12.5 - -11.2 | <0.0001 |

480 Bold print indicates statistical significance.

481 *Table 2: Effects of sensor correction on main MBW outcomes. Relative difference is calculated as 100*(corrected*
482 *- standard)/standard.*

483 **Table 3**

| | | n | Standard | | | Corrected | | | Difference | | | |
|-----------------------|------------------|----|----------|------|----------------------|-----------|------|----------------------|------------|------|-------------|-------------------|
| | | | mean | SD | p-value [#] | mean | SD | p-value [#] | mean | % | 95% CI | p-value* |
| LCI [TO] | Healthy controls | 85 | 7.12 | 0.51 | | 6.30 | 0.40 | | 0.82 | 11.5 | 0.76 - 0.88 | <0.0001 |
| | Cystic fibrosis | 62 | 9.99 | 2.21 | | 8.69 | 1.77 | | 1.30 | 13.1 | 1.14 - 1.47 | <0.0001 |
| | Difference | | -2.87 | | <0.0001 | -2.38 | | <0.0001 | | | | |
| FRC [L] | Healthy controls | 85 | 1.87 | 0.95 | | 1.73 | 0.89 | | 0.14 | 7.7 | 0.07 - 0.13 | <0.0001 |
| | Cystic fibrosis | 62 | 1.31 | 0.61 | | 1.17 | 0.53 | | 0.14 | 10.7 | 0.12 - 0.16 | <0.0001 |
| | Difference | | 0.56 | | <0.0001 | 0.56 | | <0.0001 | | | | |
| FRC [mL/kg] | Healthy controls | 85 | 45.1 | 10.5 | | 41.6 | 9.8 | | 3.51 | 7.8 | 3.33 - 3.70 | <0.0001 |
| | Cystic fibrosis | 62 | 42.3 | 8.0 | | 37.9 | 7.2 | | 4.40 | 10.4 | 4.03 - 4.76 | <0.0001 |
| | Difference | | 2.86 | | 0.0631 | 3.75 | | 0.0086 | | | | |

*paired t test; #unpaired t test

Bold print indicates statistical significance.

484

485 *Table 3: Summary of the differences in Lung Clearance Index (LCI) and functional residual capacity (FRC) between healthy controls (HC) and patients with cystic fibrosis (CF)*
 486 *in the retrospective dataset before (standard) and after (corrected) the application of the sensor correction function.*