Lung diffusing capacity for nitric oxide measured by two commercial devices: a randomised crossover comparison in healthy adults

Thomas Radtke ®[1,5], Quintin de Groot[1,2,5], Sarah R. Haile ®[3], Marion Maggi[1], Connie C.W. Hsia[4] and Holger Dressel[1]

[1]Division of Occupational and Environmental Medicine, Epidemiology, Biostatistics and Prevention Institute, University of Zurich & University Hospital Zurich, Zurich, Switzerland. [2]Zurich University of Applied Sciences, School of Health Professions, Institute of Physiotherapy, Winterthur, Switzerland. [3]Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland. [4]Dept of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, TX, USA. [5]These authors contributed equally.

Corresponding author: Thomas Radtke (thomas.radtke@uzh.ch)

Shareable abstract (@ERSpublications)
Large discrepancies between commercial devices to measure single-breath diffusing capacity of the lung for nitric oxide in healthy subjects caution against pooling or direct comparison of measurements obtained using different protocols and devices https://bit.ly/3vKyF7U

This article has supplementary material available from openres.ersjournals.com

Abstract

In Europe, two commercial devices are available to measure combined single-breath diffusing capacity of the lung for nitric oxide (D_{LNO}) and carbon monoxide (D_{LCO}) in one manoeuvre. Reference values were derived by pooling datasets from both devices, but agreement between devices has not been established.

We conducted a randomised crossover trial in 35 healthy adults (age 40.0 ± 15.5 years, 51% female) to compare D_{LNO} (primary end-point) between MasterScreen™ (Vyaire Medical, Mettawa, IL, USA) and HypAir (Medisoft, Dinant, Belgium) devices during a single visit under controlled conditions. Linear mixed models were used adjusting for device and period as fixed effects and random intercept for each participant.

Difference in D_{LNO} between HypAir and MasterScreen was 24.0 mL·min^{-1}·mmHg^{-1} (95% CI 21.7–26.3). There was no difference in D_{LCO} (-0.03 mL·min^{-1}·mmHg^{-1}, 95% CI -0.57–0.12) between devices while alveolar volume (V_A) was higher on HypAir compared to MasterScreen™ (0.48 L, 95% CI 0.45–0.52). Disparity in the estimation of V_A and the rate of NO uptake ($K_{NO}=D_{LNO}/V_A$) could explain the discrepancy in D_{LNO} between devices. Disparity in the estimation of V_A and the rate of CO uptake ($K_{CO}=D_{LCO}/V_A$) per unit of V_A offset each other resulting in negligible discrepancy in D_{LCO} between devices. Differences in methods of expiratory gas sampling and sensor specifications between devices likely explain these observations.

These findings have important implications for derivation of D_{LNO} reference values and comparison of results across studies. Until this issue is resolved, reference values, established on the respective devices, should be used for test interpretation.

Introduction

Lung diffusing capacity (DL) measures the conductance of gas transfer from alveolar air to capillary haemoglobin. The combined measurement of diffusing capacity for nitric oxide (D_{LNO}) and carbon monoxide (D_{LCO}) has recently been summarised in a technical standards document by a Task Force of the European Respiratory Society (ERS) [1]. D_{LNO} has been mainly used in research settings in healthy people and in various cardiopulmonary diseases; however, its additional value for use in clinical practice has yet to be determined. To date, two devices are commercially available in Europe: the MasterScreen™ PFT Pro, by Vyaire Medical (Mettawa, IL, USA) (hereafter referred to as MasterScreen); and the HypAir by Medisoft (Dinant, Belgium) (hereafter referred to as HypAir).

Large sets of normal values have been published for adults using prospectively collected data on the MasterScreen [2], and by pooling existing data [1] from three studies on healthy subjects [3–5] collected on MasterScreen and HypAir devices and a modified Jaeger D_{LCO} device to allow additional measurement of D_{LNO} [3]. Recent data [6] suggest substantial differences in predicted D_{LNO} values between the two reference equations by MUNKHOLM *et al.* [2] and ZAVORSKY *et al.* [1]. Different study populations, devices, testing protocols and analysis methods are likely contributing factors. Interesting observations were made by MUNKHOLM *et al.* [2] showing that the study by AGUILANIU *et al.* [4] produced the highest D_{LNO} predicted values when compared to each individual study [3, 5] contributing data to the official ERS reference equations [1]. Since AGUILANIU *et al.* [4] contributed a large dataset to the official ERS reference equation [1] (about 54% of total) and used the HypAir device, we speculated that the observed difference between the two equations [6] may be partly due to differences between devices used to measure D_{LNO}. To examine this possibility, we designed a randomised crossover study to compare the MasterScreen and HypAir devices in healthy subjects under controlled conditions and using measurement protocols recommended by the respective manufacturer. A secondary aim was to investigate the intrasession variability of both devices.

Methods

We conducted a single-centre randomised crossover trial at the University Hospital of Zurich, Switzerland between October 2019 and January 2020. Study participants were invited to the research laboratory (430 m above sea level) to perform spirometry and combined D_{LNO}–D_{LCO} measurements on the MasterScreen and HypAir devices in random order during a single study visit. Participants were advised not to perform vigorous physical activities for at least 24 h prior to testing. Coffee and meal consumption were restricted 3 h before the study visit. At the beginning of each visit, height and weight were measured to the nearest 0.1 mm and 0.1 kg, respectively. Body mass index was calculated based on height and weight.

Ethics

This study does not fall under the scope of the Human Research Act (HRA) in Switzerland. The study was designed to compare the output of two different pulmonary function devices with the results not being used for diagnostic purposes and/or to providing any treatment advice. The ethical committee of the canton of Zurich confirmed with a declaration of responsibility that ethical approval was not necessary for this study (2019-02026). All participants provided written consent to participate in this study. The study was registered with Clinicaltrials.gov (NCT04016597).

Participants

Healthy male and female adults were recruited at the Epidemiology, Biostatistics and Prevention Institute at the University of Zurich, the physiotherapy master's programme at the Zurich University of Applied Sciences (ZHAW) and the University Hospital Zurich, Switzerland. Inclusion criteria were healthy subjects aged 18 years and older. Exclusion criteria were current smoking or smoking within the last 12 months, chronic lung disease (*e.g.* asthma, COPD), a forced expiratory volume in 1 s (FEV_1) and/or forced vital capacity (FVC) below the lower limit of normal [7], acute respiratory infection, previous thoracic surgery, a body mass index >30 kg·m^{-2} and pregnancy.

Randomisation

We used simple randomisation (1:1 ratio) to allocate study participants to start the measurements either with the MasterScreen or HypAir device. A computer-generated list of random numbers was created by an independent person not involved in the study using the online randomisation tool accessible at https://www.randomizer.org. The generated list contained the numbers 1 or 2, where 1 implied the participant starts with the MasterScreen device and 2 implied the participant starts with the HypAir device. Access to the list was restricted to two independent persons not involved in this study. Allocation concealment was ensured using central randomisation, by *ad hoc* request of the allocation sequence *via* phone. This was done after the participant provided verbal consent to participate in the study and inclusion and exclusion criteria were verified. Masking of participants and outcome assessors performing pulmonary function tests was not possible in this study design.

Quality control

Prior to the start of the study, both devices underwent technical check-ups and rigorous quality control by technicians of the respective companies or service providers. Each day the devices were manually calibrated using the three-flow method and a calibrated 3-L syringe. Besides volume calibration, a gas calibration was performed using automated procedures for helium (He), carbon monoxide (CO), nitric oxide (NO) and oxygen (O_2). To ensure high-quality measurements, two members of the study team (QdG, MM) served as biological controls allowing us to detect any relevant fluctuations in D_{LNO} values

during the course of the study. Both team members (male: age 26 years and female: age 54 years, nonsmokers and free of any chronic disease impacting pulmonary function tests) performed weekly diffusing capacity measurements on both devices.

Measurement protocol

All measurements were performed on the MasterScreen™ PFT Pro (Vyaire Medical) and the HypAir (Medisoft) devices. Technical specifications for both devices are given in table S1. During spirometry and diffusing capacity measurements, participants were asked to stay seated to avoid any influence of changes in cardiac output on diffusing capacity measurements. They were allowed to drink water between test manoeuvres. Tests were done in the following order on each device: 1) slow spirometry, 2) forced spirometry, and 3) D_{LNO}–D_{LCO}. At least three technically acceptable manoeuvres were performed for both slow and forced spirometry following established standards [8]. The test with the highest value of the two best tests (*i.e.* two tests within 150-mL difference) was used in analysis. In regard to D_{LNO}, at least three technically correct tests (*e.g.* no Valsalva or Muller manoeuvre, inspired volume ⩾90% of vital capacity) were performed on each device following technical standards [9]. Additional tests (maximum of five) were done if the two best tests were not within 17 mL·min^{-1}·mmHg^{-1} [9]. In-between D_{LNO}–D_{LCO} tests, a 5-min break was allowed for complete wash-out of test gases before the new test started. After all tests had been completed on one device, a 5-min rest was ensured before starting with measurements on the other device.

Study end-points

The primary end-point was D_{LNO} (in mL·min^{-1}·mmHg^{-1}). Secondary outcomes were D_{LCO}, rate constant for NO or CO removal from alveolar gas (*i.e.* permeability factor, κNO or κCO), physiological rate of NO or CO uptake from alveolar gas (transfer coefficient of the lung for nitric oxide (K_{NO}) or carbon monoxide (K_{CO})) where K=κ/(barometric pressure–water vapour pressure) and numerically equals the corresponding DL/V_A for NO or CO, alveolar volume (V_A), change in alveolar NO fraction ($\Delta FANO$=expired−inspired NO concentration); change in alveolar CO fraction ($\Delta FACO$=expired−inspired CO concentration); breath-hold time, inspired and expired concentrations for NO, CO, He and O_2.

Statistical analyses and sample size calculation

Since there were no data available on which we could base our power calculations, we used data from our own pilot study (n=6) during which we measured team members on the two different devices. The intraclass correlation coefficient (ICC) for D_{LNO} values measured with the MasterScreen and HypAir devices was 0.96 with a 95% confidence interval (95% CI) of 0.85–1.0. Since all participants of the pilot study were experienced in performing spirometry and D_{LNO}–D_{LCO} measurements, we decided to follow a more conservative approach. Therefore, with an estimated ICC of 0.85 (95% CI 0.75–0.95), 31 participants were required (ICCest Calculation; calculated with nQuery Advisor 7.0) for primary end-point analysis. To account for possible dropout, we aimed to recruit 35 participants.

Descriptive data are presented as number (per cent) or means±SD. Diffusing capacity outcomes from HypAir and MasterScreen were analysed with a linear model adjusted for repeated measurements and reported as mean (95% confidence interval). Comparisons of primary (D_{LNO}) and secondary end-points between devices were calculated using a linear mixed model [10] adjusting for device (MasterScreen *versus* HypAir, coded as 0, 1) and period (*i.e.* 1st device used or 2nd device) as fixed effects and random intercept for each participant. Intrasession variability of HypAir and MasterScreen devices was calculated using ICC using a two-way-mixed model. Precision of D_{LNO} values was quantified by the within-subject standard deviation (SDws=root mean square error) calculated by the root mean square (RMS) method and the coefficient of variation (CV) [11, 12]. Repeatability was calculated with 1.96*1.96*$\sqrt{2}$*SDws (95% confidence interval). Intra-device repeatability was calculated as 1.96 * $\sqrt{2}$ *SDws (95% level of confidence). ICCs and their 95% confidence intervals were calculated for D_{LNO}, D_{LCO} and V_A using a two-way mixed model [consistency, single measurement (ICC, 3.1)] [13].

Results

35 participants were recruited and completed all measurements without experiencing any adverse events (figure 1). Characteristics of the study population stratified by test period are given in table 1.

Table 2 provides an overview of between-device differences in D_{LNO} (primary end-point) and all secondary end-points including all individual tests and adjusted for repeated measures. Mean raw values for D_{LNO}, D_{LCO}, V_A, κNO and κCO from all individuals' tests performed on each of the two devices are shown in figure 2. Individual mean raw data for inspired and expired gas concentrations from MasterScreen and HypAir are summarised in table S2. Individual mean raw data for breath-hold time,
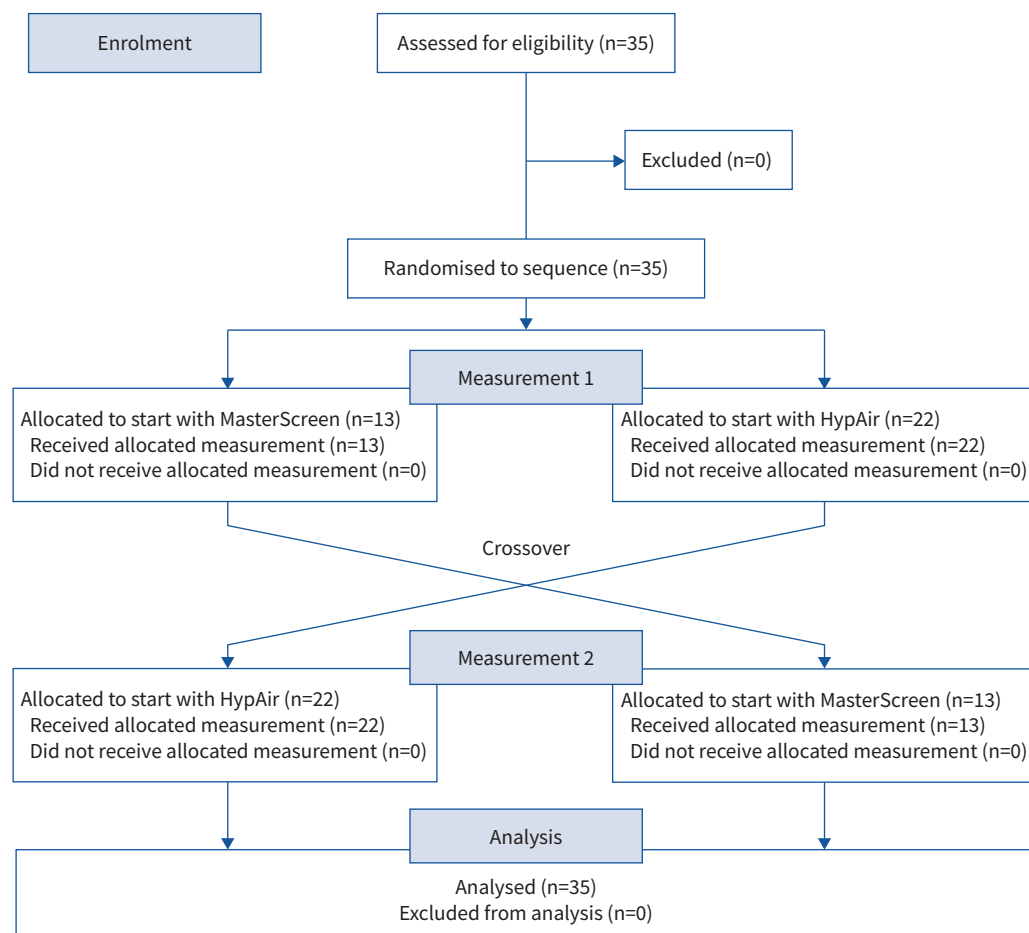
FIGURE 1 CONSORT flow chart.

inspiratory volume, K_{NO}, K_{CO}, κCO, κNO, Δ alveolar carbon monoxide fraction (FACO) and Δ alveolar nitric oxide fraction (FANO) are provided in figures S1 and S2.

In mixed linear models adjusted for period and device, the difference in D_{LNO} between HypAir and MasterScreen was 24.0 mL·min^{-1}·mmHg^{-1} (95% CI 21.7 to 26.3), see table 3. Similarly, large differences

TABLE 1 Participant characteristics stratified by test period			
Variables	**MasterScreen-HypAir**	**HypAir-MasterScreen**	**Overall**
Participants, n (%)	22 (63)	13 (37)	35 (100)
Sex, n (%)			
Male	11 (50)	6 (46)	17 (49)
Female	11 (50)	7 (54)	18 (51)
Age years	40.8±2.9	38.8±5.2	40.0±15.5
Height cm	175.5±2.5	172.8±2.7	174.5±10.1
Weight kg	73.2±11.3	63.5±12.1	69.6 ±12.4
BMI kg·m^{-2}	23.7±2.7	21.1±2.2	22.7±2.8
Heart rate beats·min^{-1}	71.6±11.2	68.5±10.9	70.5±11.0
SpO$_2$ %	97.7±0.9	97.3±1.2	97.5±1.0
FEV$_1$ MasterScreen z-score	−0.01±0.91	−0.38±0.63	−0.15±0.82
FVC MasterScreen z-score	0.55±0.92	−0.02±0.61	0.34±0.85
FEV$_1$ HypAir z-score	0.20±0.92	−0.22±0.70	0.04±0.86
FVC HypAir z-score	0.64±0.87	0.11±0.59	0.44±0.81

Data are presented as mean±SD, unless otherwise stated. BMI: body mass index; S_{pO_2}: oxygen saturation measured by pulse oximetry; FEV$_1$: forced expiratory volume in 1 s; FVC: forced vital capacity.

TABLE 2 Descriptive analysis of primary and secondary end-points between MasterScreen and HypAir

Variables	MasterScreen	HypAir	Per cent difference
Primary end-point			
D_{LNO} mL·min^{-1}·mmHg^{-1}	121.30 (110.13 to 132.48)	145.73 (132.04 to 159.41)	−16.5 (−17.8 to −15.1)
Secondary end-points			
D_{LCO} mL·min^{-1}·mmHg^{-1}	28.47 (25.81 to 31.13)	28.15 (25.46 to 30.84)	1.4 (−0.1 to 2.9)
Alveolar volume L	5.80 (5.38 to 6.21)	6.28 (5.85 to 6.71)	−8.0 (−8.9 to −7.0)
Breath-hold time s	6.28 (6.17 to 6.39)	5.93 (5.81 to 6.05)	6.2 (4.6 to 7.8)
Inspiratory volume L	4.66 (4.28 to 5.04)	4.68 (4.31 to 5.05)	−0.5 (−1.7 to 0.7)
κNO s^{-1}	−6.13 (−6.25 to −6.00)	−7.23 (−7.44 to −7.02)	−14.6 (−17.3 to −11.9)
κCO s^{-1}	−0.018 (−0.019 to −0.018)	−0.022 (−0.022 to −0.021)	−15.8 (−17.2 to −14.4)
K_{NO} mL·min^{-1}·mmHg^{-1}·L^{-1}	20.8 (19.7 to 21.9)	23.1 (21.8 to 24.4)	−9.3 (−10.4 to −8.1)
K_{CO} mL·min^{-1}·mmHg^{-1}·L^{-1}	4.9 (4.6 to 5.1)	4.5 (4.2 to 4.7)	10.2 (8.6 to 11.9)
$\Delta FANO$ ppm	−38.3 (−38.9 to −37.7)	−42.8 (−44.1 to −41.5)	−9.4 (−12.8 to −6.0)
$\Delta FACO$ %	−0.11 (−0.12 to −0.11)	−0.13 (−0.13 to −0.13)	−10.9 (−12.4 to −9.4)
NO, expired/inspired ratio	0.117 (0.107 to 0.127)	0.100 (0.090 to 0.110)	19.2 (13.9 to 24.6)
O_2, expired/inspired ratio	0.877 (0.870 to 0.884)	0.887 (0.873 to 0.900)	−1.4 (−2.7 to −0.01)
CO, expired/inspired ratio	0.470 (0.458 to 0.481)	0.475 (0.464 to 0.468)	1.3 (−0.2 to 2.8)
He, expired/inspired ratio	0.737 (0.718 to 0.756)	0.683 (0.665 to 0.702)	−10.1 (−10.9 to −9.3)

Data are presented as mean (95% CI). D_{LCO}: diffusing capacity of the lung for CO; D_{LNO}: diffusing capacity of the lung for NO; κCO: permeability factor for CO; κNO: permeability factor for NO; K_{NO}: transfer coefficient of the lung for NO; K_{CO}: transfer coefficient of the lung for CO; FACO: fractional concentration of CO in the alveolar space; FANO: fractional concentration of NO in the alveolar space; $\Delta FACO$ (expired–inspired CO concentration); $\Delta FANO$ (expired–inspired NO concentration). Ratios of expired to inspired gas concentrations are also shown.

were noticed in V_A (8%), κNO (15%) and κCO (16%), while D_{LCO} was not different between HypAir and MasterScreen (table 3).

Intrasession variability

Intrasession variability characteristics for MasterScreen and HypAir are displayed in table 4. All participants fulfilled the test quality criteria for D_{LNO} (*i.e.* the two highest tests were within 17 mL·min^{-1}·mmHg^{-1}), except one participant who performed five tests on the HypAir device without reaching the quality criterion.

Biological monitoring

During the study, two team members completed a total of 21 D_{LNO}–D_{LCO} measurements. Both individuals showed slight variation in D_{LNO} with a mean variation of 6.13 mL·min^{-1}·mmHg^{-1} (CV 3.17%) and 2.60 mL·min^{-1}·mmHg^{-1} (CV 2.72%) on the HypAir device, respectively. On the MasterScreen device, mean variation of D_{LNO} was 4.56 mL·min^{-1}·mmHg^{-1} (CV 3.07%) and 2.64 mL·min^{-1}·mmHg^{-1} (CV 3.06%), respectively. Additional diffusing capacity outcomes (*i.e.* D_{LCO}, V_A, K_{CO}, K_{NO}, κCO, and κNO) are given in table S3. These two well-trained team members showed the same systematic between-device differences as in our study population.

Discussion

This randomised crossover study was designed to directly compare D_{LNO} (primary end-point) measurements in healthy, non-smoking adults using two devices commercially available in Europe. The intrasession variability characteristics for D_{LNO} and D_{LCO} were comparable between the two devices and similar to previous studies in healthy people [9, 14] and those with chronic lung disease [15], indicating that both devices are internally consistent in measuring lung diffusing capacity outcomes. However, there are substantial differences in D_{LNO} between the MasterScreen and HypAir devices, with values on average 17% higher by HypAir than MasterScreen. In contrast, the simultaneously measured D_{LCO} was similar (1% difference) between the two devices. Published studies reporting D_{LNO} reference values also showed a similar discrepancy between HypAir and MasterScreen devices. MUNKHOLM *et al.* [2] measured D_{LNO} using the Jaeger MasterScreen Pro and obtained values significantly lower than that obtained by AGUILANIU *et al.* [4] using HypAir with an initial 14% He and selecting D_{LNO} values from the manoeuvre yielding the highest D_{LCO}. Other studies including ZAVORSKY *et al.* [5] using HypAir and 9.47% initial He concentration and VAN DER LEE *et al.* [3] using a Jaeger D_{LCO} device with substantial modifications and an added chemiluminescence NO cell, and a recent ERS Task Force document [1], reported D_{LNO} values
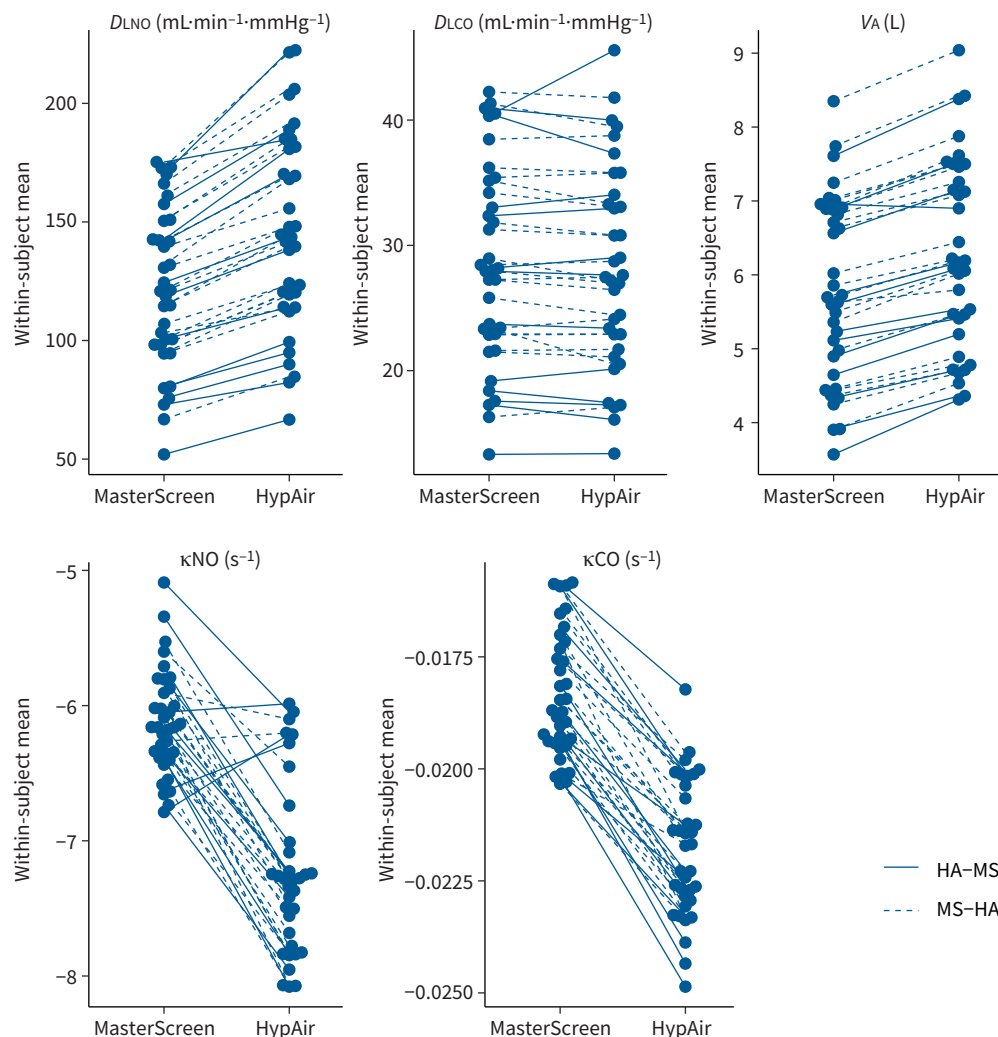
FIGURE 2 Comparison of mean individual raw data for diffusing capacity of the lung for nitric oxide (D_{LNO}), diffusing capacity of the lung for carbon monoxide (D_{LCO}), alveolar volume (V_A), and permeability factor for NO (κNO) and permeability factor for CO (κCO) between MasterScreen and HypAir in 35 subjects measured sequentially in random order (HypAir-MasterScreen (HA-MS) or MasterScreen-HypAir (MS-HA)).

intermediate between HypAir (14% initial He) and MasterScreen (~10% initial He). Despite the variations in subject populations, measurement methods and analysis, comparisons of prior studies are consistent with our current finding of a higher D_{LNO} using HypAir than using MasterScreen.

Multiple factors likely contributed to the selective discrepancy in D_{LNO} and are systematically discussed below. Between-device differences in diffusing capacity outcomes should be interpreted based on the magnitude of effects, *i.e.* statistical significance *versus* clinical/physiological relevance.

Breath-hold time

D_{LNO} and D_{LCO} decreases while V_A increases with increasing breath-hold times [2, 16], and NO was taken up from the inspired gas mixture faster than CO. In healthy adults, a difference of 2 s in breath-hold time (*i.e.* 4 s *versus* 6 s) resulted in a mean difference in D_{LNO} and V_A of ~6 mL·min^{-1}·mmHg^{-1} and <100 mL, respectively [16]. In both HypAir and MasterScreen devices, breath-hold times are calculated using the same equations [17], but they operate with different software interfaces. The small difference in breath-hold times between measurements on the two devices, −0.33 s (95% CI −0.42 to −0.24) shorter on HypAir (table 2), should not contribute significantly to the differences in D_{LNO} and V_A between the two devices.

TABLE 3 Mixed linear models on differences in D_{LNO} and secondary end-points between MasterScreen and HypAir devices

Variables	Coefficient	95% CI	p-value
Primary end-point			
D_{LNO} mL·min^{-1}·mmHg^{-1}	24.0	21.7 to 26.3	<0.0001
Secondary end-points			
D_{LCO} mL·min^{-1}·mmHg^{-1}	−0.03	−0.57 to 0.12	0.20
Alveolar volume L	0.48	0.45 to 0.52	<0.0001
Breath-hold time s	−0.33	−0.42 to −0.24	<0.0001
Inspiratory volume L	0.008	−0.028 to 0.026	0.66
κNO s^{-1}	−1.08	−1.21 to −0.94	<0.0001
κCO s^{-1}	−0.0035	−0.004 to −0.003	<0.0001
K_{NO} mL·min^{-1}·mmHg^{-1}·L^{-1}	2.2	1.9 to 2.5	<0.0001
K_{CO} mL·min^{-1}·mmHg^{-1}·L^{-1}	−0.41	−0.46 to −0.37	<0.0001
ΔFANO ppm	−4.40	−5.15 to −3.66	<0.0001
ΔFACO %	−0.015	−0.016 to −0.013	<0.0001
NO, expired/inspired ratio	−0.017	−0.012 to −0.014	<0.0001
O$_2$, expired/inspired ratio	0.012	0.005 to 0.019	0.0014
CO, expired/inspired ratio	−0.007	−0.012 to −0.002	0.0063
He, expired/inspired ratio	−0.054	−0.058 to −0.050	<0.0001

Data are presented as mean (95% confidence interval). Positive coefficients means that HypAir gave larger measurements than MasterScreen. D_{LCO}: diffusing capacity of the lung for CO; D_{LNO}: diffusing capacity of the lung for NO; κCO: permeability factor for CO; κNO: permeability factor for NO; K_{NO}: transfer coefficient of the lung for NO; K_{CO}: transfer coefficient of the lung for CO; FACO: fractional concentration of CO from the alveolar space; FANO: fractional concentration of NO from the alveolar space; ΔFACO (expired–inspired CO concentration); ΔFANO (expired–inspired NO concentration). Ratios of expired to inspired gas concentrations are also shown.

Anatomical dead space

The two devices use the same equation to calculate total dead space based on body weight, with a 50 mL difference in the apparatus dead space (table S1). However, the anatomical dead space varies with age, sex, height, body size and lung volume [18]; this factor may introduce inaccuracy in the assumed anatomical dead space but is expected to similarly affect results obtained on both devices.

He dilution

The recommended initial He concentration is 10% [1]. The initial He concentration by MasterScreen was 9.9%. The manufacturer of HypAir recommended a gas mixture containing 14% He (default setting). The thermal conductivity He analysers in the two devices have similar resolution (table S1). The accuracy of MasterScreen is ±0.05% or 2%, whichever is greater; the accuracy of HypAir is <1%, which could range from 50% lower or up to 20 times higher than MasterScreen. The response time of HypAir is 25 to 50 times slower than that by MasterScreen. He concentrations measured from a reservoir (HypAir) may be more constant than that measured in real time (MasterScreen). The starting He concentrations were also different (14% HypAir *versus* 9.9% MasterScreen). Each factor may cause minimal disparity, but cumulatively they could potentially contribute to the 10% difference in He dilution (figure S2), which in

TABLE 4 Intrasession variability characteristics for MasterScreen and HypAir

Variables	MasterScreen				HypAir			
	Measurement error#	Repeatability¶	CV %	ICC	Measurement error#	Repeatability¶	CV %	ICC
D_{LNO} mL·min^{-1}·mmHg^{-1}	6.76	18.73	4.89	0.972 (0.952–0.985)	5.97	16.52	4.05	0.962 (0.934–0.979)
D_{LCO} mL·min^{-1}·mmHg^{-1}	1.16	3.22	3.80	0.986 (0.975–0.992)	0.97	2.70	3.38	0.981 (0.968–0.990)
V_A L	0.11	0.30	1.77	0.995 (0.992–0.997)	0.11	0.31	1.79	0.993 (0.987–0.996)

Data were measured at a target breath-hold time of 5 s. CV: coefficient of variation; ICC: intraclass correlation coefficient; D_{LNO}: diffusing capacity of the lung for NO; D_{LCO}: diffusing capacity of the lung for CO; V_A: alveolar volume. #: the measurement error (or within-subject standard deviation, SDws) was calculated by the root mean square method; ¶: repeatability of gas diffusing measurements was calculated as the SDws from all three single-breath tests separately and multiplied by 2.77 (95% level of confidence).

turn could account for an 8% higher V_A estimated by HypAir relative to MasterScreen. However, any difference in V_A is expected to similarly affect the estimates of both D_{LNO} and D_{LCO}, suggesting there are additional causes for the larger discrepancy in D_{LNO} estimation.

Methods of expiratory gas sampling

With HypAir, the average inspired and expired gas concentrations are measured before and after breath-hold in separate collection bags. With MasterScreen, there is a single bag for inspiratory gas while expiratory gas concentrations are measured in real time within the device; the details of sampling and calculations are not disclosed. We were unable to extract the raw gas disappearance curves to directly verify the linearity of NO or CO uptake or the constancy of He concentration during breath-hold. Differences in the two measurement approaches may also have affected the observed rates of gas uptake (K_{NO}, K_{CO}).

Measurement of NO concentration and uptake

While average inspired NO was higher by HypAir than MasterScreen (47.6 *versus* 43.4 ppm, respectively), expired NO concentration was slightly lower (4.8 *versus* 5.1 ppm, respectively). The expired/inspired NO ratio was on average 14.5% higher by MasterScreen than HypAir; the corresponding K_{NO} was 11% higher by HypAir than MasterScreen (23.1 *versus* 20.8 mL·min^{-1}·mmHg^{-1}·L^{-1}, respectively). Since D_{LNO} is the product of $K_{NO} \times V_A$, the 11% higher K_{NO} combined with the 8% higher estimates of V_A by HypAir than MasterScreen could potentially explain the observed 20% net difference in D_{LNO} between devices.

The two devices use different models of electrochemical NO cells from the same manufacturer (CiTicel®, City Technology Ltd, Portsmouth, UK). Specifications (table S1) show that while the two sensors have comparable sensitivity and similar resolution and repeatability, the model used by HypAir (3MNT) has a larger measurement range and different accuracy (0 to 1000 ppm, relative accuracy <1%) compared to that used by MasterScreen (7NT, range 0 to 100 ppm, absolute accuracy 3 ppm). Without side-by-side performance comparison of the two NO sensors under identical controlled conditions and knowledge of the algorithms used for real time expiratory gas sampling and calculation of NO uptake by MasterScreen, it remains unclear whether the two models behave comparably at the low levels of expiratory NO concentrations (average 4–5 ppm with the lowest values <2.5 ppm) or to what extent differences in sensors contributed to the discrepancy in D_{LNO} estimates. Furthermore, sensor output signal can drift over time, thereby altering the magnitude of discrepancy both within and between devices.

Measurement of CO concentration and uptake

Both inspired and expired CO concentrations were higher by HypAir than MasterScreen (by 11% and 10%, respectively) (table 2), while K_{CO} was 9% lower by HypAir than MasterScreen. Given that V_A was 8% higher by HypAir than MasterScreen, the discrepancies in K_{CO} and V_A cancelled each other resulting in negligible net discrepancy (1%) in D_{LCO} estimates between the two devices. The electrochemical CO cells used in the two devices exhibit similar accuracy and response time with different ranges; it is unclear whether they have similar resolutions.

The NO electrochemical cell of MasterScreen has a narrower range of detection than that of HypAir, suggesting that the linearity for NO decay during breath-hold may be more strictly maintained in MasterScreen. However, the CO electrochemical cell of HypAir has a narrower range of detection than that of MasterScreen, suggesting that the linearity of CO decay may be more strictly maintained in HypAir. Without the ability to directly examine the concentration curves, in the presence of the other between-device technical differences mentioned above, and in the absence of an established gold standard, we cannot state whether one device is "more accurate" than the other.

Limitations of the study

We studied healthy subjects with a broad age range, under strictly controlled laboratory conditions, and using a randomised crossover design (*i.e.* each participant acts as its own control) to minimise confounding factors. We used simple randomisation (1:1 ratio) that resulted in an unequal number of participants starting with the HypAir device (*i.e.* 63% were randomly allocated to start with this device). Stratified block randomisation would have been the preferred randomisation method to ensure similar group sizes. Nevertheless, participant characteristics were well balanced along both sequence groups. We compared the two devices following the respective measurement conditions recommended by each manufacturer, resulting in different initial gas concentrations and breath-hold times. Strictly matching these conditions within and among subjects and/or by the use of a simulation system may allow pinpointing the source(s) of the observed discrepancy. We did not study subjects with lung disease, although it is unlikely that the systematic differences in D_{LNO} between the two devices will disappear when testing subjects with

cardiorespiratory diseases; for example, in obstructive airway disease ventilatory inhomogeneity plus a short breath-hold time may further accentuate any discrepancy in V_A, K_{NO} and K_{CO} between devices.

Conclusions

The rapid pulmonary uptake of NO relative to that of CO [1] increases the susceptibility of D_{LNO} to methodological variations such as in the initial and final gas concentrations, breath-hold time, and device-related differences in the quantification of helium dilution and NO and CO uptake rates. Disparity in the estimation of V_A and K_{NO} could fully explain the observed discrepancy in D_{LNO} measurement between devices. Disparity in the estimation of V_A and K_{CO} offset each other resulting in negligible discrepancy in D_{LCO} measurement between devices. These large disparities measured in healthy subjects have important implications for the derivation of D_{LNO} reference values and the comparison of results across studies, which in turn impact the utility of D_{LNO} as a biomarker of lung disease. Further studies will examine whether and how the presence of lung disease alters device-related disparity in D_{LNO} measurements. Given these uncertainties and the need to avoid systematic errors, caution must be exercised regarding the pooling or comparison of D_{LNO} measurements obtained using different protocols and devices.

References

1   Zavorsky GS, Hsia CCW, Hughes JMB, *et al.* Standardisation and application of the single-breath determination of nitric oxide uptake in the lung. *Eur Respir J* 2017; 49: 1600962.

2   Munkholm M, Marott JL, Bjerre-Kristensen L, *et al.* Reference equations for pulmonary diffusing capacity of carbon monoxide and nitric oxide in adult Caucasians. *Eur Respir J* 2018; 52: 1500677.

3   van der Lee I, Zanen P, Stigter N, *et al.* Diffusing capacity for nitric oxide: reference values and dependence on alveolar volume. *Respir Med* 2007; 101: 1579–1584.

4   Aguilaniu B, Maitre J, Gienet S, *et al.* European reference equations for CO and NO lung transfer. *Eur Respir J* 2008; 31: 1091–1097.

5   Zavorsky GS, Cao J, Murias JM. Reference values of pulmonary diffusing capacity for nitric oxide in an adult population. *Nitric Oxide-Biol Chem* 2008; 18: 70–79.

6   Radtke T, Hua-Huy T, Dressel H, *et al.* Of the need to reconcile discrepancies between two different reference equations for combined single-breath D_{LNO}–D_{LCO} in systemic sclerosis. *Eur Respir J* 2019; 53: 1802109.

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

8   Miller MR. Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.

9   Zavorsky GS, Murias JM. A small amount of inhaled nitric oxide does not increase lung diffusing capacity. *Eur Respir J* 2006; 27: 1251–1257.

10   Stephen SS. Cross-Over Trials in Clinical Research (Statistics in Practice), 2nd Edn. Chichester, John Wiley & Sons, Ltd., 2003.

11   Zavorsky GS. The rise in carboxyhemoglobin from repeated pulmonary diffusing capacity tests. *Respir Physiol Neurobiol* 2013; 186: 103–108.

12   Gluer CC, Blake G, Lu Y, *et al.* Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. *Osteoporos Int* 1995; 5: 262–270.

13   Bland JM, Altman DG. Measurement error. *BMJ* 1996; 312: 1654.

14   Desjardin A, Creveuil C, Bergot E, *et al.* Assessment of concordance between diffusion of carbon monoxide through the lung using the 10 s breath-hold method, and the simultaneous NO/CO technique, in healthy participants. *Respir Physiol Neurobiol* 2020; 273: 103319.

15   Radtke T, Benden C, Maggi-Beba M, *et al.* Intra-session and inter-session variability of nitric oxide pulmonary diffusing capacity in adults with cystic fibrosis. *Respir Physiol Neurobiol* 2017; 246: 33–38.

16   Dressel H, Filser L, Fischer R, *et al.* Lung diffusing capacity for nitric oxide and carbon monoxide: dependence on breath-hold time. *Chest* 2008; 133: 1149–1154.

17   Jones RS, Meade F. A theoretical and experimental analysis of anomalies in the estimation of pulmonary diffusing capacity by the single breath method. *Q J Exp Physiol Cogn Med Sci* 1961; 46: 131–143.

18   Anthonisen NR, Fleetham JA. Ventilation: total, alveolar, and dead space. *Comprehensive Physiology* 2011: 113–129.