Long-Term Intersession Variability for Single-Breath Diffusing Capacity

Matthew J. Hegewald^a Robert L. Jensen^a John G. Teeter^b Robert A. Wise^c
Richard J. Riese^b Richard D. England^b Richard C. Ahrens^d Robert O. Crapo^a
Neil R. MacIntyre^e

^aDivision of Pulmonary and Critical Care Medicine, University of Utah School of Medicine and Intermountain Medical Center, Salt Lake City/Murray, Utah, ^bPfizer Global Research and Development, New London, Conn.,

^cDivision of Pulmonary and Critical Care Medicine, The Johns Hopkins University School of Medicine, Baltimore, Md.,

^dUniversity of Iowa School of Medicine, Iowa City, Iowa, and ^eDuke University Medical School, Durham, N.C., USA

Key Words

Pulmonary diffusing capacity • Pulmonary function tests • Intersession variability • Biological control

Abstract

Background: Characterizing long-term diffusing capacity (DL_{CO}) variability is important in assessing quality control for DL_{CO} equipment and patient management. Long-term DL_{CO} variability has not been reported. **Objectives:** It was the aim of this study to characterize long-term variability of DL_{CO} in a cohort of biocontrols and to compare different methods of selecting a target value. **Methods:** Longitudinal DL_{CO} monitoring of biocontrols was performed as part of the inhaled insulin development program; 288 biocontrols were tested twice monthly for up to 5 years using a standardized technique. Variability, expressed either as percent change or DL_{CO} units, was assessed using three different target values. **Results:** The 90th percentile for mean intersession change in DL_{CO} was between 10.9 and 15.8% (2.6–4.1 units) depending on the target value. Variability was lowest when the mean of all DL_{CO} tests was used as the target value and highest when

the baseline DL_{CO} was used. The average of the first six DL_{CO} tests provided an accurate estimate of the mean DL_{CO} value. Using this target, the 90th percentile for mean intersession change was 12.3% and 3.0 units. Variability was stable over time and there were no meaningful associations between variability and demographic factors. **Conclusions:** DL_{CO} biocontrol deviations >12% or >3.0 units, from the average of the first six tests, indicate that the instrument is not within quality control limits and should be carefully evaluated before further patient testing.

Copyright © 2011 S. Karger AG, Basel

Introduction

The single-breath diffusing capacity for carbon monoxide (DL_{CO}) is an important and widely used test that provides a quantitative measure of gas transfer in the lungs [1–8]. Defining variability in lung function is important for both laboratory quality control and for patients with lung disease. Variability associated with longitudinal measurements of DL_{CO} can be attributed to both technical (in-

strument) and biological factors [2, 9]. Quantifying the background intersession variability is essential to determine when changes in DL_{CO} exceed the normal biological and technical variability in the measurement (the 'noise'). Changes in DL_{CO} that exceed this normal background variability may be attributed to changes due to the disease, therapy or exposure (the 'signal'). The American Thoracic Society and European Respiratory Society (ATS/ERS) guidelines recommend longitudinal testing of biological controls as a quality control measure [10]. However, the best method of establishing a target value and acceptable range of deviation has not been well established. Knowing DL_{CO} variability over time can help determine if changes in a biocontrol are due to the normal variability inherent in the test or to changes in the instrument that needs to be addressed. Defining intra-individual variability for DL_{CO} is clinically important. Serial measurements of DL_{CO} are commonly performed to assess disease activity in response to therapeutic interventions [11–13] and to monitor the effects of drug or toxin exposures [14–16].

Long-term variability for DL_{CO} has not been studied longitudinally for more than 1 year [17, 18]. Further, the ATS/ERS recommendation that a change in DL_{CO} of >10% over 1 year is clinically significant is not founded in high quality studies [17].

We analyzed serial DL_{CO} measurements from a large cohort of biocontrols that underwent longitudinal DL_{CO} monitoring as part of the Exubera™ (insulin human inhalation powder; Pfizer, Inc., New York, N.Y., and Nektar Therapeutics, San Carlos, Calif., USA) developmental program [19] to assess: (1) long-term DL_{CO} variability in healthy controls; (2) whether intersession variability is more stable across the range of measured DL_{CO} values in terms of absolute DL_{CO} units (ml CO/min/mm Hg) or percent change; (3) factors influencing DL_{CO} variability, and (4) whether there is a learning effect with DL_{CO} measurement.

Methods

Study Participants

Biological controls were technical staff working in clinics participating in inhaled human insulin (Exubera) clinical studies. Biocontrols performed serial DL_{CO} measurements for up to 5 years as part of the quality control program. Participants were free from active lung disease based on self-reporting; detailed health assessment was not performed. Smoking status was not reported in all biocontrols. Current and past smokers were not excluded. Retrospective analysis of the biological control dataset was performed. Institutional review board approval was waived for biocontrol testing.

Monitoring and Measurements

Study personnel participated in a 2-day group training session at a central location and were required to pass a written test and exhibit test performance proficiency on five volunteers to be certified to test patients. Study personnel were primarily nurses and medical assistants working in endocrinology offices; certified pulmonary function technicians were not used. All study sites used the same pulmonary function testing equipment (Collins CPL; nSpire Health, Longmont, Colo., USA). All lung function equipment was factory tested using a DL_{CO} simulator [18] and required to be accurate within 2.0 DL_{CO} units. Subsequent DL_{CO} simulator testing was not performed. Instruments were maintained according to the manufacturer's specifications and were calibrated and leak tested daily. If calibration or leak testing failed, the manufacturer guided identification and repair of the problem. Standardized protocol-specific software provided real-time quality-control prompts and feedback to technicians. Offsite independent quality assessment was performed by a clinical research organization (nSpire Health) on all tests to emulate the review done by a pulmonary function laboratory medical director.

The biocontrol testing protocol suggested test performance at least bimonthly. The goal of each testing session was to obtain two acceptable and repeatable efforts. Acceptable efforts were defined using ATS criteria [20]: inspiratory vital capacity $\geq 90\%$ of the largest forced vital capacity; inspiratory time ≤ 2 s; breath-hold time of 9–11 s; and an expiratory time ≤ 4 s. Discard volumes and alveolar sampling volumes were set at 750 and 500 ml, respectively. The intrasession repeatability goal was 3.0 DL_{CO} units [20]. DL_{CO} units were measured as ml CO/min/mm Hg. The result reported from a test session was the mean of two acceptable efforts. Test data were transmitted to the reading center within 24 h for quality review and feedback to the technician. Only sessions with two acceptable tests varying by $\leq 15\%$ were used in the present analysis. To avoid bias in estimating DL_{CO} variability, no threshold value for biological control intersession variability was set. Hemoglobin and carboxyhemoglobin concentrations were not obtained as we felt it was not reasonable to perform repeated venipuncture in biological controls.

Statistical Analysis

The database was restricted to biological controls with at least six measurements made in a period of at least 6 months and no more than 60 months. Outliers, defined as tests with ≥ 11 DL_{CO} units difference from the initial value, were excluded from analysis by inspection of a histogram showing that 99.2% of the tests were normally distributed at ± 11 DL_{CO} units. Data of each subject were aligned as number of months from the first test performed that met inclusion criteria. Data were compared to a target value in three ways:

(1) The baseline-adjusted DL_{CO} determined by dividing each DL_{CO} measurement for an individual by the initial measurement, expressed as percent:

$$\text{Baseline adjusted } DL_{CO} = \frac{[DL_{CO} \text{ at time } t_{(0-60)} / \text{baseline } DL_{CO}]}{\times 100}$$

(2) The mean adjusted DL_{CO} determined by taking the average of all DL_{CO} measurements for an individual and dividing each DL_{CO} measurement by this mean value, expressed as percent:

$$\text{Mean adjusted } DL_{CO} = \frac{[DL_{CO} \text{ at time } t_{(0-60)} / \text{mean } DL_{CO}]}{\times 100}$$

Fig. 1. Data selection.

(3) The average of the first six DL_{CO} measurements (X_6) for an individual and dividing each DL_{CO} measurement by this mean value, expressed as percent:

$$X_6 \text{ adjusted } DL_{CO} = [DL_{CO} \text{ at time } t_{(0-60)} / X_6 DL_{CO}] \times 100$$

Variability was estimated as the standard deviation of the residuals from linear regression of time versus DL_{CO} (normalized) and expressed as percent. Variability was also estimated in terms of absolute DL_{CO} units using the difference from the baseline, mean adjusted and average of the first six values.

To determine the number of tests needed to obtain an acceptable estimate of the mean of all DL_{CO} measurements for a biocontrol, the baseline value and the average of the first 2–10 measurements were subtracted from the mean of all tests.

Table 1. Baseline demographic characteristics

Sex	Female	245 (85)
	Male	43 (15)
Race	White	219 (76.0)
	Black	15 (5.2)
	Hispanic	33 (11.5)
	Asian	16 (5.6)
	Other	5 (1.7)
Age, years	Mean $\pm$ SD	39.3 $\pm$ 10.4
	Range	18 – 70
Height	Mean $\pm$ SD	165.6 $\pm$ 9.2
	Range	130 – 210
Weight	Mean $\pm$ SD	70.7 $\pm$ 16.2
	Range	45 – 133
Baseline DL_{CO} , ml CO/min/mm Hg	Mean $\pm$ SD	23.5 $\pm$ 5.5
	Range	13.6 – 47.3

Figures in parentheses are percentages.

To analyze the ‘learning effect’, DL_{CO} variability was calculated as the standard deviation of the measured DL_{CO} – the mean DL_{CO} expressed in DL_{CO} units for the first 12 months compared with the standard deviation calculated over years 2–5. An F test was used to determine if there was a significant difference in variability for year 1 compared with years 2–5.

To determine if DL_{CO} variability was significantly different when expressed in DL_{CO} units or percent change across the range of DL_{CO} values, variability was calculated in each quartile of baseline DL_{CO} value using DL_{CO} units and percent change.

To investigate factors influencing DL_{CO} variability, variability was analyzed by multiple regression models that included demographic factors, spirometry measurements and baseline and mean DL_{CO} values.

Results

Data Selection and Subject Characteristics

A total of 288 biological controls in 162 sites had DL_{CO} tests that met the study criteria. Data selection resulting in the final database used for analysis is listed in figure 1. 14,129 DL_{CO} tests were completed. The final data set consisted of 8,585 DL_{CO} tests; 2,120 DL_{CO} tests did not meet the time criteria inclusive of 6 months to 5 years, 2,466 tests (17% of total) were excluded due to lack of centralized quality review, and 958 tests did not meet intrasession repeatability criteria (within test session repeatability $\leq 15\%$). The average number of acceptable DL_{CO} tests per biocontrol was 29.8 ± 14.0 (range 6–80). Baseline demographic characteristics are listed in table 1. The bio-

Fig. 2. a Variation in DL_{CO} adjusted to baseline over 5 years. **b** Difference from baseline DL_{CO} by year.

Fig. 3. a Variation in DL_{CO} adjusted to the mean of the first six measurements over 5 years. **b** Difference from the mean of the first six measurements of DL_{CO} by year.

controls were predominately female (85%) and white (76%); the mean age was 39.3 years. Smoking status was determined in a sample of 81 biocontrols: 5% current, 27% past and 68% never.

Intersession DL_{CO} Variability

The 90th percentile for intersession variability, calculated using the mean adjusted DL_{CO} as the target for each individual test, expressed as a percentage change from the mean value was 10.9%, or 2.6 DL_{CO} units. The inter-

session variability using the baseline DL_{CO} as the reference expressed as a percentage change and in DL_{CO} units by year are shown in figure 2. The 90th percentile for intersession variability expressed as a percentage change from the baseline test was 15.8%, or 4.1 DL_{CO} units. The intersession variability using the average of the first six DL_{CO} measurements as the reference for subsequent tests expressed as a percentage change and in DL_{CO} units by year are shown in figure 3. The 90th percentile for intersession variability expressed as a percentage change com-

Fig. 4. Average absolute difference between the mean of the first to 10th DL_{CO} measurement and the overall mean.

pared with the average of the initial six DL_{CO} measurements was 12.3%, or 3.0 DL_{CO} units. Intersession DL_{CO} variability is summarized in table 2.

The average of the first six DL_{CO} tests deviated from the mean of all DL_{CO} tests by 0.69 DL_{CO} units; averaging additional tests did not significantly improve the accuracy at estimating the mean DL_{CO} (fig. 4).

Variability was stable over time. DL_{CO} variability did not change significantly for subsequent years compared to the first year. Variability expressed as the standard deviation determined using the mean DL_{CO} value as the target for subsequent tests was 1.49 DL_{CO} units for year 1 and 1.51 DL_{CO} units for years 2–5 ($F = 1.397$; $p = 0.237$). The stability of DL_{CO} variability over time indicates there was no significant ‘learning effect’ in the first year of testing.

DL_{CO} Variability Expressed in DL_{CO} Units versus Percent Change

Significant differences in DL_{CO} variability were observed across the quartiles of baseline DL_{CO} values for variability expressed in terms of DL_{CO} units and percent change from the mean value. Variability expressed in terms of percent change from the mean value was higher in the lowest quartile, and variability measured as absolute change in DL_{CO} units from the mean value was larger in the highest quartile. Neither measurement was ideal across the range of DL_{CO} values.

Table 2. Long-term intersession variability of DL_{CO} using three different methods of selecting the target value

	Intersession variability (90th percentile)	
	%	DL _{CO} units
Initial measurement	15.84	4.14
Average of first six measurements	12.25	2.98
Average of all measurements	10.91	2.62

Factors Influencing DL_{CO} Variability

Multiple regressions were performed with DL_{CO} variability expressed in DL_{CO} units as the dependent variable against demographic data, spirometry values (forced vital capacity, forced expiratory volume in 1 s) and baseline and mean DL_{CO} values as the independent variable. Some significant associations were observed ($p < 0.05$); however, all r^2 values were < 0.010 , indicating that $< 1.0\%$ of the variation was explained from these variables. The association between sex and DL_{CO} variability was not significant ($p = 0.08$). The association between age and DL_{CO} variability was significant ($p < 0.05$), with an r^2 value of 0.0029.

Discussion

Intersession DL_{CO} variability was assessed in biological controls tested with a standardized technique over a 5-year period. The main finding is that the 90th percentile for mean intersession change in DL_{CO} for up to 5 years was stable. Depending on the method of selecting the target value, intersession variability ranged between 10.9 and 15.8% (2.6–4.1 DL_{CO} units). Intersession variability was lowest when the mean DL_{CO}, the average of all measurements, was used to compare each individual DL_{CO} test and highest when the initial DL_{CO} was used as the reference for subsequent tests. While the mean DL_{CO} provides the best representation of the actual DL_{CO} over time of a biological control, it requires the calculation of the average of all previous DL_{CO} tests and is not practical for routine clinical use. A reasonable alternative to using the overall mean DL_{CO} is to use an average of the first six DL_{CO} tests to compare subsequent test values. The average of the first six DL_{CO} tests provides an accurate estimate of the overall mean DL_{CO} and averaging additional tests improves the accuracy very little. A DL_{CO} test result

that deviates by $>12.3\%$ or >3.0 units from the average of the first six DL_{CO} tests exceeds the 90th percentile and should prompt further actions such as equipment recalibration, maintenance, retesting with another biological control or testing with a DL_{CO} simulator [18].

Long-term intersession variability for DL_{CO} has not been well defined. Previous data are not robust. ATS/ERS guidelines presently recommend that a change of $>10\%$ over 1 year in healthy adults is significant [17]. There is no specific recommendation on what baseline values should be used for comparison. The ATS/ERS recommendation is based on one study by Hathaway et al. [21] who measured DL_{CO} five times over a 1-year period in 8 healthy pulmonary laboratory technicians. For the average subject tested using the two best functioning machines, intra-individual variability was 10% (95% confidence interval). Jensen et al. [9] measured DL_{CO} three times over 6 months in 11 healthy subjects using multiple instruments and found in the well-performing instruments a significant change in DL_{CO} was >3.0 DL_{CO} units (95% confidence interval). Drummond et al. [22] measured intersession DL_{CO} variability in nonsmoking diabetics with no significant lung disease over 6 months. For the 699 subjects tested using a highly standardized technique, the 90th percentile for intersession variability was 11.3% and 3.1 units. Variability was significantly higher for the 948 subjects studied using routine clinical testing techniques (90th percentile for intersession variability of 20.0% and 5.4 units). Our study used a highly standardized technique similar to the one used in the study of Drummond et al. [22]. We found similar intersession variability that extended up to 5 years (90th percentile for intersession variability of 12.3% and 3.0 units when referenced to the average of the first six DL_{CO} tests). Studies addressing intersession are slightly higher than the current ATS/ERS guideline recommendation when standardized techniques are employed. The current study is the only one to address DL_{CO} variability beyond 1 year; a change of $>12\%$ or >3.0 DL_{CO} units in a biological control suggests a technical issue needs to be addressed, assuming there is no physiological change in the biocontrol.

Intersession variability can be expressed in terms of percent change or absolute change in DL_{CO} units. Expressing variability as a percentage assumes that variability is greater for higher DL_{CO} values (i.e. non-homoscedastic). We found that neither the percent change nor the absolute change in DL_{CO} units is optimal for comparing subsequent DL_{CO} tests across the full range of measurements. Other investigators have found an increase in

variability within a test session for lower DL_{CO} values using the percentage difference [23]. For reasons of simplicity, we propose using an absolute difference of 3.0 DL_{CO} units between the target values, calculated from the average of the first six DL_{CO} tests, as the threshold for deviation that would prompt further action across the range of DL_{CO} values. However, for DL_{CO} values <20.0 units (corresponding to the lowest quartile for our data), a change of 2.0 DL_{CO} units will exceed the 90th percentile and could be considered as the threshold value.

A limitation of the present study is that the majority of biological controls were female and Caucasian. It is not known if variability would be different in a more heterogeneous population although we found no difference in variability for men compared with women. Age did have a significant effect on variability although the effect was small ($r^2 = 0.0029$). Hemoglobin and carboxyhemoglobin concentrations were not measured, and therefore, no adjustment of the DL_{CO} was made as suggested by the ATS/ERS guidelines [10]. The failure to adjust for hemoglobin will potentially increase our measured intersession variability because of the known association of measured DL_{CO} with the menstrual cycle and the preponderance of females in the study [24]. Smoking status was not obtained from all biocontrols preventing analysis of DL_{CO} variability based on smoking status. However, in the subset of biocontrols in which smoking status was queried, only 5% were active smokers. The small number of active smokers will minimize the effect of not adjusting for carboxyhemoglobin. The biological controls reported no overt lung disease but their respiratory health status was not fully characterized throughout the study interval. Biocontrols with mild, stable asthma and chronic obstructive pulmonary disease (COPD) were not excluded. However, Drummond et al. [22] found no change in intersession variability when diabetics with mild stable asthma or COPD were excluded from his analysis, suggesting that including biocontrols with mild, stable obstructive lung disease will not have a significant impact on measured variability. Not all biocontrols had an equal number of measurements (range 6–80), which might act as a selection bias. Our inclusion criteria required test sessions to have intra-session repeatability of $\leq 15\%$. This is greater than the current standard of $\leq 10\%$ and may have increased our estimate of intersession variability [10]. A third-party quality review was performed for all included tests, resulting in the exclusion of 17% of tests. These tests were excluded to assure that the quality of the tests was equal to those done in a hospital pulmonary laboratory. The same model of pulmonary function testing equip-

ment (Collins CPL; nSpire Health) was used in all sites. Instrument accuracy and reproducibility can vary significantly between different models of instruments, and our results may not apply to other DL_{CO} instruments [9].

Extrapolation of our results to patients with lung disease should be done with caution. Longitudinal variability for DL_{CO} in patients with chronic lung disease has not been well studied but is important as clinical decisions are made based on changes in DL_{CO} over time. Pulmonary function test variability is generally greater in patients with lung disease than in patients without lung disease [17]. Spirometry is more variable over time in patients with chronic lung disease than in normal subjects [17, 25, 26]. In one study of COPD patients tested over 1 week, a change in DL_{CO} of >4.0 units was considered significant [27]. As mentioned previously, a change in DL_{CO} of >3.1 units over a period of 6 months, using a highly standardized technique, in diabetics without lung disease was considered significant [22]. The usual clinical practice is to compare a baseline DL_{CO} with a subsequent DL_{CO} measurement after treatment or observation to assess disease progression [3, 7, 8]. Our results in healthy biocontrols indicate that a change of approximately 15% or 4.0 DL_{CO} units $\geq$ 6 months after an initial DL_{CO} test is the upper limit of normal variability (90th percentile). It is reasonable to suggest that DL_{CO} variability is higher in patients with chronic lung disease than what we have reported in generally healthy subjects. In patients with

chronic lung disease, a change of at least 15% or 4.0 DL_{CO} units compared with a baseline value is likely needed to represent a significant change. More studies with longitudinal measurement of DL_{CO} in patients with lung disease are needed.

This is the first study to examine long-term (>1 year) DL_{CO} variability. DL_{CO} variability in healthy biocontrols tested for up to 5 years was approximately 2.6–4.1 DL_{CO} units (90th percentile) depending on the target value used for comparison. Routine biological control testing is an essential component for laboratory quality control. For biological controls, we suggest averaging the initial six DL_{CO} tests and then using this baseline value to compare all subsequent measurements. A change in DL_{CO} of >3.0 units compared with this baseline suggests that the instrument is not within quality control limits.

Acknowledgment

The data discussed in this article were obtained from clinical trials funded by Pfizer Global Research and Development.

Financial Disclosure and Conflict of Interest

None of the authors have reported potential conflicts of interest with any companies or organizations whose products or services may have been discussed in this article.

References

- 1 Hughes JM: The single breath transfer factor (TL_{CO}) and the transfer coefficient (KCO): a window onto the pulmonary microcirculation. *Clin Physiol Funct Imaging* 2003;23:63–71.
- 2 Crapo RO, Jensen RL, Wanger JS: Single-breath carbon monoxide diffusing capacity. *Clin Chest Med* 2001;22:637–649.
- 3 Martinez FJ, Flaherty K: Pulmonary function testing in idiopathic interstitial pneumonias. *Proc Am Thorac Soc* 2006;3:315–321.
- 4 Sun XG, Hansen JE, Oudiz RJ, Wasserman K: Pulmonary function in primary pulmonary hypertension. *J Am Coll Cardiol* 2003;41:1028–1035.
- 5 Aduen JF, Zisman DA, Mobin SI, Venegas C, Alvarez F, Biewend M, et al: Retrospective study of pulmonary function tests in patients presenting with isolated reduction in single-breath diffusion capacity: implications for the diagnosis of combined obstructive and restrictive lung disease. *Mayo Clin Proc* 2007;82:48–54.
- 6 Cotton DJ, Soparkar GR, Grahon BL: Diffusing capacity in the clinical assessment of chronic airflow limitation. *Med Clin North Am* 1996;80:549–564.
- 7 Takeda S, Funakoshi Y, Kadota Y, Koma M, Maeda H, Kawamura S, et al: Fall in diffusing capacity associated with induction therapy for lung cancer: a predictor of postoperative complication? *Ann Thorac Surg* 2006;82:232–236.
- 8 Wasfi YS, Rose CS, Murphy JR, Silveira LJ, Grutters JC, Inoue Y, et al: A new tool to assess sarcoidosis severity. *Chest* 2006;129:1234–1245.
- 9 Jensen RL, Teeter JG, England RD, Howell HM, White HJ, Pickering EH, et al: Sources of long-term variability in measurements of lung function: implications for interpretation and clinical trial design. *Chest* 2007;132:396–402.
- 10 MacIntyre N, Crapo RO, Viegi G, Johnson DC, van der Grinten CP, Brusasco V, et al: Standardisation of the single-breath determination of carbon monoxide uptake in the lung. *Eur Respir J* 2005;26:720–735.
- 11 Flaherty KR, Andrei AC, Murray S, Fraley C, Colby TV, Travis WD, et al: Idiopathic pulmonary fibrosis: prognostic value of changes in physiology and six-minute-walk test. *Am J Respir Crit Care Med* 2006;174:803–809.
- 12 Sietsma K: Sarcoidosis and the diffusing capacity for carbon monoxide. *Sarcoidosis* 1990;7:12–14.
- 13 Stam H, Splinter TA, Versprille A: Evaluation of diffusing capacity in patients with a restrictive lung disease. *Chest* 2000;117:752–757.
- 14 Camus P, Martin WJ 2nd, Rosenow EC 3rd: Amiodarone pulmonary toxicity. *Clin Chest Med* 2004;25:65–75.
- 15 Ceglia L, Lau J, Pittas AG: Meta-analysis: efficacy and safety of inhaled insulin therapy in adults with diabetes mellitus. *Ann Intern Med* 2006;145:665–675.

- 16 Dimopoulou I, Galani H, Dafni U, Samakovi A, Roussos C, Dimopoulos MA: A prospective study of pulmonary function in patients treated with paclitaxel and carboplatin. *Cancer* 2002;94:452–458.
- 17 Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, et al: Interpretative strategies for lung function tests. *Eur Respir J* 2005;26:948–968.
- 18 Jensen RL, Crapo RO: Diffusing capacity: how to get it right. *Respir Care* 2003;48:777–782.
- 19 Wise RA, Teeter JG, Jensen RL, England RD, Schwartz PF, Giles DR, et al: Standardization of the single-breath diffusing capacity in a multicenter clinical trial. *Chest* 2007;132:1191–1197.
- 20 American Thoracic Society: Single-breath carbon monoxide diffusing capacity (transfer factor). Recommendations for a standard technique – 1995 update. *Am J Respir Crit Care Med* 1995;152:2185–2198.
- 21 Hathaway EH, Tashkin DP, Simmons MS: Intraindividual variability in serial measurements of DL_{CO} and alveolar volume over one year in eight healthy subjects using three independent measuring systems. *Am Rev Respir Dis* 1989;140:1818–1822.
- 22 Drummond MB, Schwartz PF, Duggan WT, Teeter JG, Riese RJ, Ahrens RC, et al: Inter-session variability in single-breath diffusing capacity in diabetics without overt lung disease. *Am J Respir Crit Care Med* 2008;178:225–232.
- 23 Punjabi NM, Shade D, Patel AM, Wise RA: Measurement variability in single-breath diffusing capacity of the lung. *Chest* 2003;123:1082–1089.
- 24 Sansores RH, Abboud RT, Kennell C, Haynes N: The effect of menstruation on the pulmonary carbon monoxide diffusing capacity. *Am J Respir Crit Care Med* 1995;152:381–384.
- 25 Herpel LB, Kanner RE, Lee SM, Fessler HE, Sciurba FC, Connett JE, et al: Variability of spirometry in chronic obstructive pulmonary disease: results from two clinical trials. *Am J Respir Crit Care Med* 2006;173:1106–1113.
- 26 Pennock BE, Rogers RM, McCaffree DR: Changes in measured spirometric indices. What is significant? *Chest* 1981;80:97–99.
- 27 Robson AG, Innes JA: Short term variability of single breath carbon monoxide transfer factor. *Thorax* 2001;56:358–361.