Volumetric capnography and chronic obstructive pulmonary disease staging

Pablo V Romero[1]
Benigno Rodriguez[1]
Daniela de Oliveira[1]
L Blanch[2]
Federico Manresa[1]

[1]Laboratorio de Función Pulmonar, Servicio de Neumologia, Hospital Universitario de Bellvitge, IDIBELL, Hospitalet de Llobregat, Barcelona, Spain; [2]Critical Care Center, Hospital de Sabadell, Institut Universitari Fundació Parc Taulí, Sabadell, Spain

Abstract: Spirometry is difficult for some COPD patient to perform. Volumetric capnography could be a second choice test to evaluate the severity of functional disturbances. The aim of this work is to test this hypothesis. A total number of 98 subjects were classified either as normal ex-smokers (N = 14) or COPD patients. The latter were staged following GOLD recommendations. Spirometry and volumetric capnography recordings were obtained from each patient. Spirometry parameters, Bohr Dead Space (V_D^{Bohr}), Airways Dead Space from the pre-interface expirate corrected curve (V_D^{aw}), Phase III slope (Sl_{III}) and Volume of alveolar ejection (V_{AE}) were measured. Index of Ventilatory Efficiency (IVE), and Index of Airways Heterogeneity (IAH) were calculated as: $IVE = V_{AE}/(V_T - V_D^{aw})$ and $IAH = 1 - [(V_T - V_D^{Bohr})/(V_T - V_D^{aw})]$. In ANOCOVA analysis IAH showed the greatest association with stage (F > 40), with no significant covariant dependence on V_T. A receiver operating characteristics curve analysis showed values of the area under the curve greater than 0.9 for IAH and IVE at all stage levels, with a sensitivity = specificity value greater than 80%. We conclude that IAH and IVE can be used when spirometry cannot be reliably performed, as an alternative test to evaluate the degree of functional involvement in COPD patients.

Keywords: capnography, chronic obstructive pulmonary disease, dead space, pulmonary heterogeneity, receiver operating characteristics curve, spirometry

Introduction

Stable chronic obstructive pulmonary disease (COPD) can be evaluated by means of standard spirometry in a large majority of patients. In fact, spirometry is the gold standard to diagnose and classify COPD patients, and the severity of the disease. When properly performed, spirometry is the most reproducible, standardized and objective way of measuring flow limitation. A correct spirometry manoeuvre, however, requires the ability to clearly understand the manoeuvre, the intellectual capacity to perform it, and the absence of clinical limitations which might prevent several maximal forced expiratory manoeuvres from being performed. It is relatively common to deal with subjects with disabled intellectual capacity, with language difficulties, or clinical situations that make the forced expiratory manoeuvres painful or may even preclude them (chest pain, coronary vascular disease, minimal efforts dyspnoea, etc) In such cases spirometry is a poor cost/benefit procedure. Other situations such as intubation or tracheotomy, or patients lying in bed also limit the application of standard spirometry.

Pulmonary heterogeneity is, together with airways obstruction, a cardinal feature in the functional impairment of COPD. Heterogeneity, mostly dependent of peripheral involvement, is seen to increase with the severity of the disease; therefore, volumetric capnography, a technique that basically explores regional distribution, can be expected to correlate with the severity of disease in COPD patients. Previous authors have outlined the good relationship between changes in volumetric capnography parameters and the severity of COPD (Kars et al 1995; Kars et al 1997; Wilschut et al 1999;

and Koulouris et al 2001); however, the diagnostic value of volumetric capnography has not been systematically explored. Recent studies reduce the intrinsic variability of the parameters, either by using internal relationships (Romero 2001), a procedure that reduces the noise-to-signal ratio, or by using the reference values reported by Åström et al (2000) in normal people.

Volumetric capnography parameters are measured in continuous scales. Their use as a diagnostic test requires the selection of a cut-off value in the original scale in order to define positive and negative test outcome. Comparison of the dichotomised test results with the clinical status of the patient, evaluated as the disease stage from Global Initiative for Chronic Lung Disease (GOLD) classification (a spirometry based criterion) (GOLD 2006), allows estimation of the test's diagnostic sensitivity (Se) and specificity (Sp), and determines the cut-off values that concurrently optimise Se and Sp. According to the hypothesis that volumetric capnography could be an alternative test to determine the amount of functional involvement in COPD patients, we have designed a study to compare the diagnostic value of volumetric capnography with spirometry as gold standard.

Methods

Subjects

A group of 84 patients (18 women, 66 men; age 60 ± 13 years, mean $\pm$ standard deviation) with clinical symptoms of COPD have been selected from patients attending the Pulmonary Function Laboratory. This protocol was approved by the Hospital's Research Committee. Patients were asked to give their consent for statistical use of the results of functional studies. The selection criteria were: clinical diagnosis of COPD based on clinical symptoms and history of exposure to cigarette smoking. Age limits were 30 to 80 years, and body mass index (BMI) was limited to 35. Exclusion criteria were cardiovascular comorbidity disease, sequelae of tuberculosis on the chest X-ray or any other coexistent lung disease, and neuromuscular or other diseases causing loss of ventilatory capacity. In order to exclude asthmatics, patients with a change of more than 12% in FEV_1 after bronchodilator were excluded. Subjects were classified according to the Global Initiative for Chronic Lung Disease (GOLD) criteria (GOLD 2006) into five groups:

– Group 0 (N = 33) or subjects at risk, with normal spirometry (FEV%FEC $\geq$70%, FEV_1%pred $\geq$80%), but symptoms (cough, sputum).
– Group I (N = 1) or mild COPD, with FEV_1%FVC <70%, but FEV_1%pred <80%.

– Group II (N = 20) or moderate COPD, with FEV_1%FVC < 70%, and 50% $\leq FEV_1$%pred <80%.
– Group III (N = 19) or moderate to severe COPD, with FEV_1%FVC <70%, and 30% $\leq FEV_1$%pred <50%.
– Group IV (N = 18) or severe COPD, with FEV_1%FVC <70%, and FEV_1%pred <30%.

In these five groups FEV_1 and FVC post-bronchodilator in percent of predicted were considered. A group of 14 normal subjects (6 women, 8 men; age 48 ± 11 years, mean $\pm$ SD), ex-smokers (more than 10 years of cessation) but free of symptoms was also studied, and is called "Ex-Smok" Group. The same exclusion criteria were applied. Table 1 shows the descriptive and spirometric data for the different groups. All COPD subjects were in a stable situation, with no changes in respiratory symptoms during the month prior to the study, and no acute episodes of respiratory failure more than six months. Patients were asked not to have any inhaled therapy in the previous twelve hours. They sat down for at least half an hour before the procedure.

Instrumentation

The equipment used was a Cosmo Plus System (Novametrix Medical Systems Inc, Wallinford, CT, USA), which measures CO_2 concentration by means of a mainstream infrared cell, and airflow by an integrated pitot type pneumotachograph. Optical CO_2 calibration and zero flow were performed before each measurement. The instrumental dead space (mouthpiece plus pneumotachograph and infrared cell chamber) was 18 ml. Pneumotachograph linearity was confirmed in the experimental range of flows, with increasing static flows measured by means of a rotameter (Collins, Braintree,

Table 1 Descriptive statistical data of the different groups of subjects studied

Group	N	Sex ♀/♂	Age yrs	BMI	FEV₁ %pr	FVC %pr	MEF₅₀ %pr
Ex-Smok	14	6/8	48 ± 11.1	24.2 ± 2.04	100.3 ± 8.9	93.2 ± 7.78	109.3 ± 14.7
GOLD 0	26	13/13	52 ± 14.8	27.1 ± 4.37	86.3 ± 16.2	84.2 ± 17.3	67.8 ± 18.0
GOLD I	1	0/1	57	32	88.6	98.5	42.9
GOLD II	20	2/18	63 ± 11.4	26.5 ± 4.16	58.4 ± 7.32	73.5 ± 13.0	26.1 ± 8.2
GOLD III	19	3/16	64 ± 13.1	25.5 ± 4.5	40.1 ± 5.73	59.2 ± 10.1	16.2 ± 7.8
GOLD IV	18	0/18	62 ± 9.5	25.8 ± 4.49	22.4 ± 5.2	42.4 ± 14.9	7.42 ± 2.72

Abbreviations: BMI body mass index; FEV_1%pr, forced expiratory volume at the first second in percent of predicted; FVC%pr, forced vital capacity in percent of predicted; MEF_{50}%pr, maximal expiratory flow at 50% FVC in percent of predicted. **Note:** Mean $\pm$ Standard Deviation.

Mass, USA). Expiratory flow and volume are expressed in BTPS conditions. Response characteristics of mainstream capnograph, as well as the validity of optical calibration, was determined by flowing a gas mixture containing 5% (in) or 0% (out) CO_2 in air by means of a syringe at step changing flow rates in the range of the physiological rates. Delay between flow and capnographic signal and capnograph time constant were determined by analyzing the system's step response between 5% and 0% CO_2 in air. In all cases delay was measured as the time elapsed between the beginning of flow transition according to the pneumotachographic recording, and the breakpoint of the capnographic recording. Average delay was 58.9 ± 16.2 ms. Time constant was determined as the time to reach 63.2% of the total step response, and was 24.4 ± 2.35 ms. Capnographic signal was time shifted according to the time delay.

Measurements: Capnographic recording was always performed before spirometry. Subjects were sitting during the recording procedure, breathing through a mouthpiece and with a nose clip. After two to three minutes of stable breath, CO_2 and flow signals were continuously recorded for two more minutes, digitalised at sampling frequency of 100 Hz and stored in magnetic media. Forced spirometry was performed with patient in standing position by means of a Fleisch pneumotachograph (GS-System, Collins, Braintree, MA, USA) according to ATS/ERS Standardization of Spirometry (ATS/ERS, http://www.thoracic.org/sections/publications/statements/pages/pfet/pft2.html, update 2006). Spirometry was processed by GS software, and results expressed in BTPS conditions. To accept spirometry results, at least three manoeuvres with a reproducibility of 5% and a difference between the best and second best FEV_1 and FVC below 150 ml had to be recorded. A maximum of eight forced expiratory manoeuvres were performed. Maximal values of forced expiratory flow in one second (FEV_1), Forced Vital Capacity (FVC) and Maximal Mid Expiratory Flow between 25 and 75% of FVC (MMEF) were retained. The minimum time between manoeuvres was three minutes (of longer if requested by the patient). Spirometric manoeuvres were repeated 15 minutes after inhalation of 400 μg of salbutamol by means of a spacing chamber. According to recorded observations, time consumption by spirometric tests oscillated between about 20 minutes (the fastest) and 50 minutes (the slowest) plus the time consumed by the reversibility test. Predicted spirometric values were calculated from age, height and sex according to Roca et al (1986).

Volumetric capnography parameters: A recording containing a minimum of fifteen complete breathing cycles,

at a minimal tidal volume of 150 ml was processed for every patient. Every cycle was identified from the pneumotachographic signal, isolated and analysed separately by a computer procedure. Cycles not accepted for measurements were those with a $P_{ET}CO_2$ 5 mmHg lower or higher than the average $P_{ET}CO_2$. Also discarded were cycles with a tidal volume above or below 25% of average V_T. Final results give the average of the measurements done on seven to fifteen cycles for each patient.

End tidal CO_2 fractional concentration ($F_{ET}CO_2$): to avoid the artefact of cardiac oscillations, $F_{ET}CO_2$ was measured on the CO_2 elimination (VCO_2) versus expired volume (V) curve [$VCO_2 = \int F_ECO_2 \cdot dV$], as the linear slope between expired volume and VCO_2 of the segment defined by the end expiratory 10% of the total number of expiratory samples of each breath (Figure 1). End-tidal partial pressure of CO_2 ($P_{ET}CO_2$) was calculated from barometric pressure (P_B) as: $F_{ET}CO_2 \cdot (P_B - P_{H2O})$.

Bohr Dead Space ($V_D{}^{Bohr}$) was calculated according to classical principles, assuming that end-tidal fraction of CO_2 represents alveolar fractional concentration in the equation:

$$V_D{}^{Bohr}/V_T = 1 - (V_ECO_2/V_T)/F_{ET}CO_2 \qquad [1]$$

where V_ECO_2 is the CO_2 eliminated by breath.

Pre-interface expirate (PIE) was calculated according to Wolff and Brunner's method (Wolff and Brunner 1984). In short, the volume expired when CO_2 has reached half of the end-expiratory value is doubled, the rest of the points are excluded from further analysis. Then the first derivative of the signal (dF_ECO_2/dV) is plotted against the volume, and the mean of this normalized distribution function is calculated. The volume at which this mean value is obtained represents the minimal mean volume of the convective airways, also called the pre-interface expirate or PIE (Wolff and Brunner 1984). Physiologically speaking, PIE represents the expired volume at which the interface between airways and alveolar gas is appreciated at airways opening.

Phase III slope (Sl_{III}): the curve between PIE and V_T is divided into four segments. According to Åström et al (2000), Sl_{III} was calculated as the slope of the linear regression between F_ECO_2 and Volume for the two central segments.

Series dead space ($V_D{}^{ser}$): The curve was corrected for the Sl_{III} between PIE and end tidal volume. From the corrected new curve, serial dead space volume ($V_D{}^{ser}$) was calculated by the equal area method. This procedure prevents overcorrection of phase II (Åström et al 2000). Airways dead
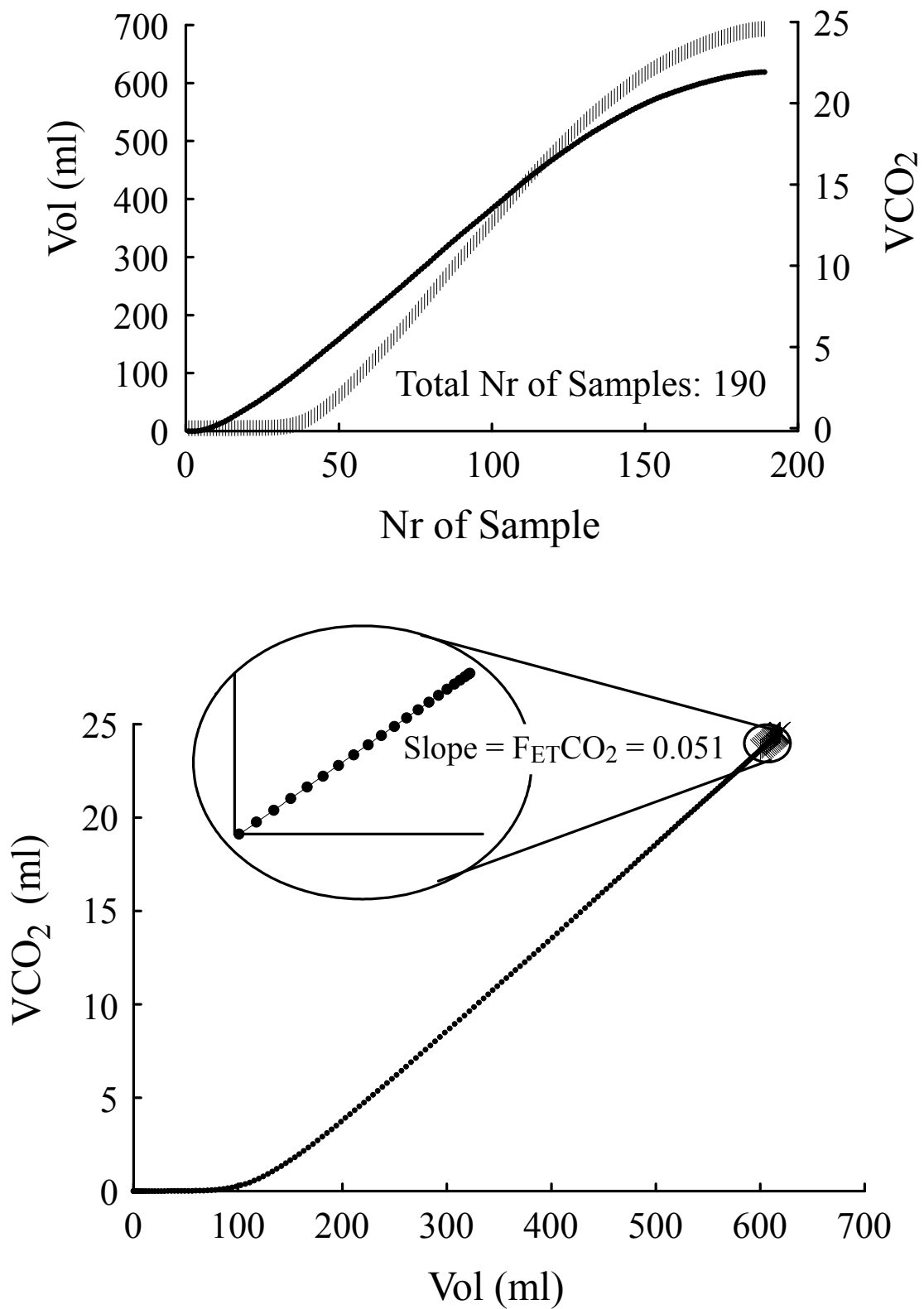
Figure 1 Top: Expired volume and VCO₂ sampled curves. Bottom: Measurements of end expiratory CO₂ concentration. For detail please see the text.

space (V_D^{aw}) was obtained by subtracting Instrumental Dead Space (18 ml) from V_D^{ser}.

Index of Alveolar Inhomogeneity (IAH): According to previous studies (Kars et al 1995; Romero 2001) the difference between V_D^{ser} and V_D^{Bohr} is mainly due to impairment of regional distribution, which distorts the curve beyond PIE. An index of heterogeneity was calculated by relating both magnitudes according to:

$$IAH(\%) = \left[1 - \frac{\left(V_T - V_D^{Bohr} \right)}{\left(V_T - V_D^{ser} \right)} \right] \times 100 \qquad [2]$$

Index of Ventilatory Efficiency (IVE): Alveolar exhaled volume or V_{AE} was determined from the VCO_2 versus expired volume or $VCO_2(V)$ curve, according to previous studies (Romero et al 1997; Blanch et al 1999). In short, after the linear fitting of the last end-expiratory segment previously described, the slope of this curve was decreased by 6% by

pivoting on the end expiratory point. The new straight line crosses the $VCO_2(V)$ curve at a single point. The volume between this and the end-expiratory point corresponds to V_{AE} (Figure 2). This volume tends to decrease as serial contamination of alveolar gas, heterogeneity, and phase II increase. As V_{AE} directly depends on V_T and V_D^{ser}, it seemed appropriate to express it in relation to these magnitudes. An index of ventilatory efficiency (IVE) has been calculated as follows:

$$IVE(\%) = \frac{V_{AE}}{V_T - V_D^{ser}} \times 100 \qquad [3]$$

We used a 6% change in end tidal slope of $VCO_2(V)$ curve rather than the 5% used in previous studies in ventilated patients, because we have observed less variability in spontaneously breathing patients when using 6% (CVar% ± SD: 9.14 ± 3.34, and 10.98 ± 4.16 at 6% and 5% respectively. N = 16 patients, twenty consecutive cycles each).
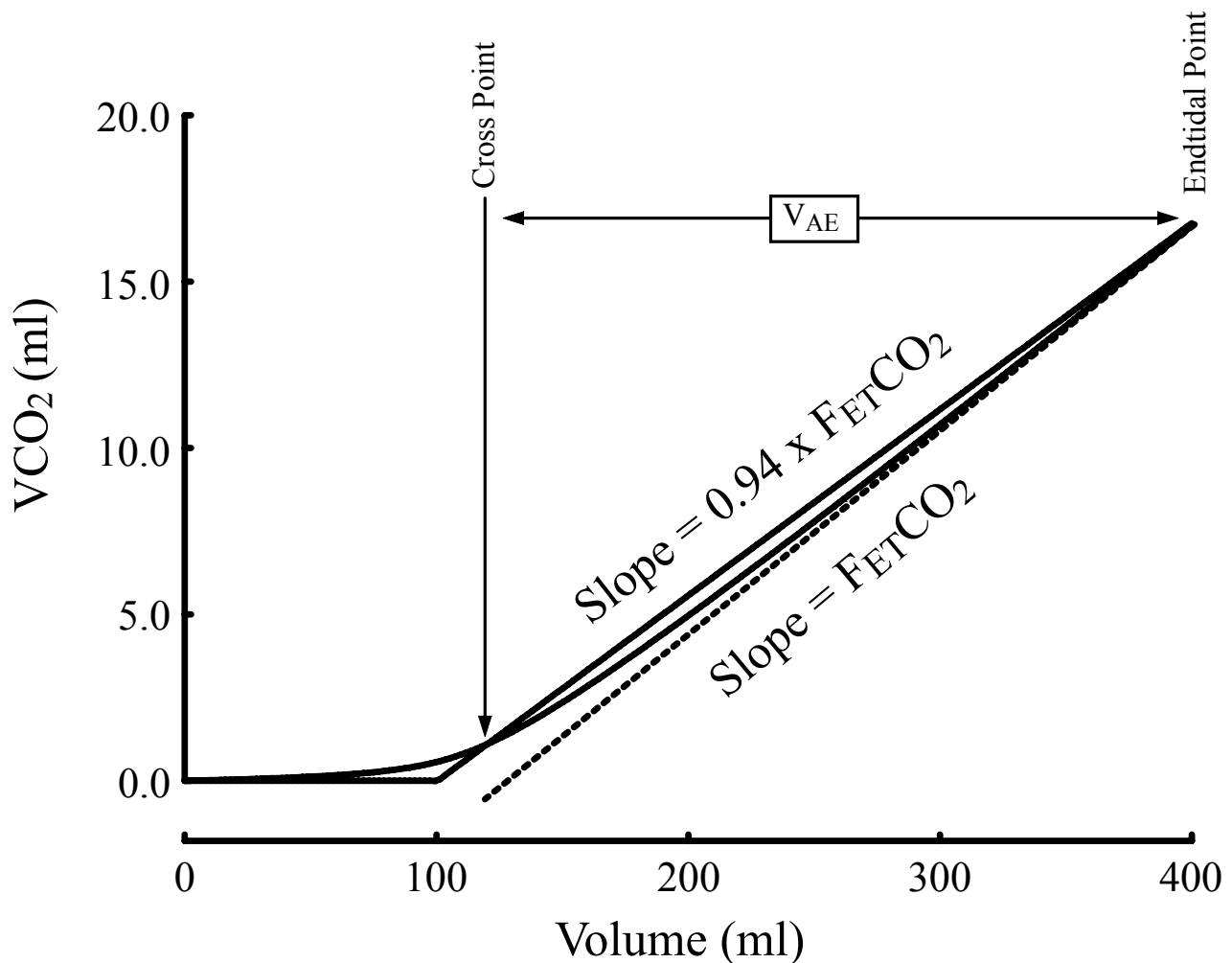
Figure 2 Measurement of the volume of alveolar ejection (VAE) from the VCO_2 versus expired volume curve. For detailed instructions, please see the text.

Statistical analysis: tests of normality (Kolmogorof-Smirnof) and homogeneity of variances (Levene homocedasticity test) were applied to data prior to deciding which test to use to compare the groups identified according to GOLD classification. For variables normally distributed, and with homogeneous variances, ANOVA analysis was performed. Tidal volume was added as covariant when necessary. If normality and homocedasticity tests failed at an error of 5%, the Kruskal-Walis rank test was applied to evaluate the differences between groups.

A multivariate model including Height, Body Mass Index, age, FEV_1%pred, Tidal volume and MMEF%pred was used to determine the association of capnography parameters with functional data, and the possible confounding effect of factors relatively independent of respiratory functional abnormalities.

Receiver operating characteristic (ROC) curve analysis was performed to test the ability of each volumetric capnography parameter to detect the different levels of disease characterised by the GOLD Stages. The analysis was performed by cutting-off the population at increasing GOLD Stages of the disease. For each Stage Level, each parameter, considered as one-class classifier, was progressively increased or decreased in step values (threshold changes). Sensitivity (Se) and specificity (Sp) were calculated for each threshold change and expressed as a percentage. For each parameter at every stage ROC curves were constructed by plotting sensitivity against the false positive rate (100-specificity). The area under the curve (AUC) and confidence intervals at 95% were calculated to determine the degree and limits of accuracy of the curves, allowing paired comparisons. To determine cut-off points for each parameter at every stage, we calculated the point at which sensitivity equals specificity (Se = Sp). Positive and Negative Predicted Values were calculated using Bayes theorem (Greiner and Gardner 2000) based on COPD prevalence observed by Kornmann et al (2003) at stages levels II, III and IV, and internal prevalence at stage 0. MedCalc® and SPSS® statistical packages were used for statistical analysis.

Results

Table 2 shows the descriptive statistics (mean, standard deviation) of capnography parameters in the COPD groups, as well as statistical data from ANOVA test. Variables ΔSl_{III} and IVE did not meet the criteria of homocedasticity, and the Kruskal-Wallis test was applied. Group I was excluded, because it had only one subject. Although no significant between groups difference was observed for

V_T, this parameter was used as covariant in the ANOVA analysis because of the documented tidal dependence of capnography variables (Åström et al 2000; Kornmann et al 2003), and the high dispersion of V_T. In Table II values of F (or χ^2) and P are given for between groups difference ($F(\chi^2)_{group}$, and P_{group}) and V_T covariant effect (F_{cov}, and P_{cov}). V_D^{aw} showed no significant between-groups difference, but did show high dependence on V_T. $P_{ET}CO_2$ showed mainly group, but also some V_T dependence. V_D^{Bohr}/V_T showed both group and tidal volume dependence. When corrected for predicted value, however, (Δ = real-predicted, $\Delta V_D^{Bohr}/V_T$) the covariant effect disappeared, but the group effect was maintained. Finally, IAH showed strong group but no V_T dependence.

Table 3 shows the results of the multivariate linear regression. Significance P gives the probability associated to the rejection of the null hypothesis stating that the regression coefficient (Coeff) for a given independent variable is different from zero, at an alpha error below 5%. Variables in the equation (P < 0.05) are marked in bold. The Beta statistic or regression coefficient for typified variables, allows a non dimensional comparison of the degree of association between dependent and independent variables. The narrowest association between FEV_1%pred and capnography parameters was observed for IAH. This association was not confounded by other associations with the independent variables considered. In contrast, $\Delta V_D^{Bohr}/V_T$ showed some extra dependence on age, but its association with FEV_1%pred was still high. ΔSl_{III} also showed a good correlation with FEV_1%pred, but was greatly confounded by a high dependence on V_T. Finally IVE was well defined by a model including FEV_1%pred, MMEF%pred, and age.

Using capnography parameters as independent variables, FEV_1% pred was well defined (R = 0.836) by a multivariate linear model including IAH and $\Delta V_D^{Bohr}/V_T$ (beta = –0.976, and 0.200 respectively) with the equation:

$$FEV_1(\%\,pred) = 112.08 - 5.60 \times IAH$$
$$+ 1.10 \times \Delta(V_D^{Bohr}/V_T)$$

Figure 3 shows the ROC curves constructed for 0, II, III and IV threshold stages of the disease at increasing values of IAH, Sl_{III}, and $\Delta V_D^{Bohr}/V_T$, or decreasing values of IVE. Table 4 shows the numerical analysis of the ROC curves, including AUC with 95% confidence intervals, the cut-off value at which sensitivity equals specificity, the value of sensitivity (or specificity) at this cut-off, positive predicted

Table 2 Comparison of groups by ANOVA or Kruskal-Wallis test. VT is used as covariant in ANOVA test. Values of χ^2 are marked with an asterisk

Group	N	STAT	V_T ml	$V_D{}^{aw}$ ml	$P_{ET}CO_2$ mmHg	IAH %	IVE %	$V_D{}^{Bohr}/V_T$ %	$\Delta(V_D{}^{Bohr}/V_T)$ %	ΔSI_{III} mmHg/L
Ex-smok	14	mean	548	115	38.2	4.62	84.8	32.0	7.48	12.78
		± SD	± 209	± 34	± 3.76	± 1.46	± 8.16	± 6.27	± 4.25	± 7.14
GOLD 0	26	mean	530	99	34.2	7.54	70.0	31.6	7.40	17.18
		± SD	± 176	± 23	± 5.37	± 2.81	± 13.9	± 4.50	± 4.02	± 9.05
GOLD II	20	mean	697	126	31.4	11.9	51.5	33.2	10.45	19.35
		± SD	± 227	± 38	± 3.31	± 3.30	± 12.1	± 5.94	± 4.39	± 11.57
GOLD III	19	mean	572	107	33.5	13.5	48.0	36.6	11.82	32.84
		± SD	± 247	± 37	± 6.1	± 2.57	± 6.97	± 6.72	± 4.35	± 26.9
GOLD IV	18	mean	558	117	37.4	17.5	39.9	40.9	15.43	40.46
		± SD	± 205	± 39	± 7.08	± 3.23	± 6.5	± 6.76	± 5.33	± 44.5
ANOCOVA		$F(\chi^2)_{group}$	1.76	1.54	3.299	45.79	64.45*	12.28	10.07	36.59*
		P_{group}	0.128	0.185	0.009	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
		F_{cov}	--	116.6	4.722	0.761	--	63.08	4.40	--
		P_{cov}	--	<0.0001	0.032	0.209	--	<0.0001	0.039	--

Abbreviations: V_T, tidal volume; $V_D{}^{aw}$, airways dead space; $P_{ET}CO_2$, end-tidal CO_2 partial pressure; IAH, index of alveolar heterogeneity; IVE, index of ventilatory efficiency; $V_D{}^{Bohr}/V_T$, Bohr dead space ratio; $\Delta(V_D{}^{Bohr}/VT)$, $V_D{}^{Bohr}/V_T$ ratio real minus predicted; ΔSI_{III}, Slope of phase III real minus predicted.
Note: F_{corr} and P_{corr} are values corrected for V_T effect in ANOVA analysis.

value (PPV), and negative predicted value (NPV) calculated either from internal prevalence at stage level 0, or from the cumulative prevalence calculated from Kornmann et al (2003) for the remaining stage levels. A visual inspection of the ROC curves in Figure 3 shows that IAH and IVE seem more accurate than Sl_{III}, and $\Delta V_D{}^{Bohr}/V_T$ to stage COPD patients. More precise information about the accuracy of capnography indices is obtained from numerical analysis of ROC curves. AUC, taken as a global summary statistic of diagnostic accuracy, shows higher values for IAH and

IVE than Sl_{III}, and $\Delta V_D{}^{Bohr}/V_T$, especially for diagnosis at low stage levels of the disease. Paired comparison between AUC showed the following differences with IAH: Sl_{III}, and $\Delta V_D{}^{Bohr}/V_T$ at stages 0 and II; $\Delta V_D{}^{Bohr}/V_T$ and IVE at stage III; and $\Delta V_D{}^{Bohr}/V_T$ at stage IV. Sensitivity (and specificity) was similar and higher than 80% at all threshold levels for IAH, and IVE. In all cases PPV decreased with the threshold level of the disease, and the opposite was observed for NPV. The positive likelihood ratio (PLR) (ratio between the probability of observing the test result in diseased individuals, and

Table 3 Results of the multivariant analysis. Variables in the model are in bold

Parameter	Height T (Sig T) beta	BMI T (Sig T) beta	Age T (Sig T) beta	VT T (Sig T) beta	FEV_1%pred T (Sig T) beta	MEF%pred T (Sig T) beta	R(multi)
IVE	−1.40 (0.164) −0.084	1.85 (0.067) 0.110	**−2.41 (0.018)** **−0.159**	−0.70 (0.48) −0.042	**2.78 (0.006)** **0.379**	2.82 (0.006) 0.379	**0.825**
IHA	1.38 (0.170) 0.079	−1.07 (0.29) −0.061	1.40 (0.16) 0.090	−0.31 (0.76) −0.017	**−14.4 (<0.0001)** **−0.826**	−1.66 (0.099) −0.218	**0.826**
$\Delta V_D{}^{Bohr}/VT$	0.050 (0.96) 0.004	−0.46 (0.64) −0.041	**2.15 (0.034)** **0.205**	−1.74 (0.08) −0.146	**−4.57 (0.0003)** **−0.436**	0.80 (0.42) 0.158	**0.560**
ΔSI_{III}	1.87 (0.06) 0.155	0.87 (0.38) 0.066	0.084 (0.93) 0.007	**−6.44 (<0.0001)** **−0.483**	**−6.33 (<0.0001)** **−0.475**	0.63 (0.53) 0.112	**0.683**
$V_D{}^{aw}$	1.54 (0.13) 0.115	−0.078 (0.94) −0.005	1.46 (0.15) 0.098	**11.06 (<0.0001)** **−0.130**	0.078 (0.94) 0.005	0.519 (0.13) 0.035	**0.748**

Abbreviations: V_T, Tidal volume; $V_D{}^{aw}$, airways dead space; IAH, index of alveolar heterogeneity; IVE, index of ventilatory efficiency; $\Delta(V_D{}^{Bohr}/V_T)$, $V_D{}^{Bohr}/V_T$ ratio real minus predicted; ΔSI_{III}, slope of phase III real minus predicted.

Table 4 Receiver operating characteristics analysis (internal pevalence)

	Stage (cut-off)	AUC (95% C.I.)	Cut-off (95% C.I.) sens = spec	PPV (range)	NPV (range)
IAH	Stage II	0.908 (0.884–0.982)	9.76 (9.43–10.1)	91.1 (89.5–92.6)	82.9 (82.1–83.3)
	Stage III	0.908 (0.833–0.957)	11.9 (11.1–12.7)	70.7 (69.6–75.8)	86.2 (81.5–92.3)
	Stage IV	0.926 (0.855–0.969)	14.2 (13.6–14.9)	53.6 (53.3–59.0)	95.8 (94.7–97.1)
IVE	Stage II	0.939 (0.873–0.977)	57.2 (55.6–58.7)	89.5 (88.2–92.3)	83.3 (80.0–85.6)
	Stage III	0.879 (0.798–0.936)	50.6 (49.9–51.3)	68.2 (67.7–69.4)	84.5 (67.6–69.1)
	Stage IV	0.902 (0.826–0.952)	45.5 (44.2–46.7)	51.7 (50.0–51.9)	95.7 (94.5–97.0)
$\Delta V_D^{Bohr}/VT$	Stage II	0.788 (0.694–0.864)	9.6 (8.54–10.74)	76.0 (72.0–78.7)	59.8 (62.2–59.6)
	Stage III	0.777 (0.683–0.855)	11.0 (9.26–12.8)	52.3 (51.8–65.5)	74.5 (73.9–81.4)
	Stage IV	0.816 (0.725–0.886)	12.4 (10.7–14.1)	35.1 (34.0–42.0)	90.6 (88.3–96.2)
ΔSI_{III}	Stage II	0.753 (0.656–0.834)	18.0 (16.9–19.1)	74.7 (72.7–74.9)	58.1 (56.0–59.1)
	Stage III	0.841 (0.753–0.906)	21.1 (20.5–21.7)	63.5 (62.9–66.7)	83.3 (84.1–83.0)
	Stage IV	0.871 (0.788–0.930)	28.4 (25.5–31.3)	85.2 (75.0–82.5)	80.6 (76.3–80.9)

the probability of observing the same result in non-diseased individuals) was globally high in IAH, followed by IVE, and $\Delta SLIII$ but only at stage IV.

Discussion

According to our results, volumetric capnography parameters seem to be a good tool to determine the degree of functional involvement in COPD patients. It is generally considered that volumetric capnography reflects the degree of pulmonary regional homogeneity in relation to gas exchange (Wolff and Brunner 1984; Fletcher and Jonson 1984; Hoffbrand 1966). From this point of view it seems reasonable to expect that volumetric capnography will reflect the progression of alveolar heterogeneity in parallel with that of the disease. Previous studies have shown some relationships between functional involvement in obstructive lung disease and specific indices obtained either from time based capnography (Toulou 1966; Smidt 1976; Worth 1986; You et al 1992) or from volumetric capnography in humans (Kars et al 1995; Kars et al 1997) and horses (Herholz et al 2002). However, no systematic analysis of the diagnostic capabilities of volumetric capnography in COPD patients has been carried on until now in humans. Recent technical and conceptual advances have allowed anticipation of better perspectives for clinical applications of capnography.

On the volumetric capnography parameters

We have used classical (V_D^{Bohr}, Sl_{III}) and newly developed parameters (VAE, PIE, V_D^{ser}), as well as combinations of both (IAH, IVE). Historically, the clinical use of volumetric capnography has been limited for several reasons, one of them being the lack of a reliable measurement of V_D^{aw}. The over-correction of phase II induces a phase III-dependent error in the classical, and most popular, equal area method.

Nowadays, we can avoid this problem by taking into account the interface expired (Kars et al 1997; Åström et al 2000; Romero et al 1997). In fact, the finding that V_D^{aw}, calculated according to Wolff and Brunner (1984), is unrelated to the disease stage in COPD supports this hypothesis.

Another factor that has limited the clinical use of volumetric capnography is the tidal dependence of the more usual capnography parameters. The use of predicted values means tidal dependence can be smoothed for V_D^{Bohr}/V_T, by subtracting the predicted value. A better result is obtained by using a common mode rejection ratio: IAH to completely eliminates the tidal dependence of V_D^{Bohr} by referring it to $V_T-V_D^{aw}$ (Equation 2). IVE, based on the alveolar ejection volume (V_{AE}) measured on the $VCO_2(V)$ curve, also uses $V_T-V_D^{aw}$ as reference eliminate the tidal dependence of V_{AE}. Both IAH and IVE parameters are related to functional troubles expressed during the alveolar phase of the capnogram. They are, therefore, physiologically related to regional distribution impairment as has been suggested by previous studies on normal subjects and COPD patients (Wilschut et al 1999; Hoffbrand 1966); however, the strong tidal dependence of Sl_{III} could not be corrected even after subtraction of the predicted value calculated from tidal volume.

In 1997 Kars et al observed a good correlation between the degree of airways obstruction and the V_D^{Bohr} normalised at 1 L V_T (R = –0.66). Our results are in concordance with this observation, as $\Delta V_D^{Bohr}/V_T$ has shown a significant correlation with $FEV_1\%$ pred. In the multivariate model we have used, however, age is significantly correlated with $\Delta V_D^{Bohr}/V_T$, a fact that was not noticed by Kars (1997) in COPD patients, or by Åström (2000) in normal subjects. In the study by Kars, only simple regressions were performed. Perhaps the age dependence we have observed is a surrogate dependence on
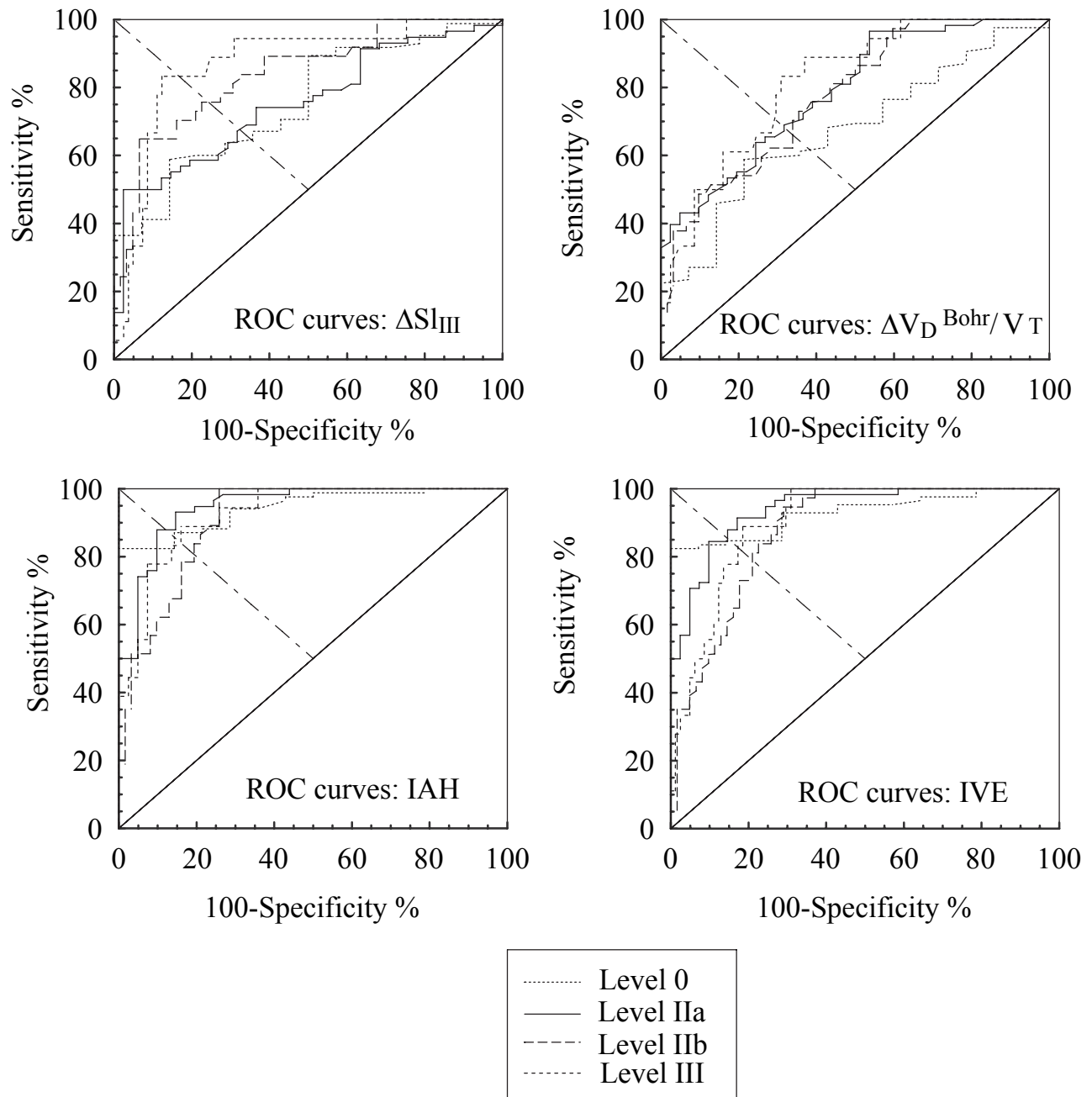
Figure 3 Receiver operating characteristics curve for each capnography parameter are plotted in separate boxes (ΔV_D^{Bohr}: Bohr dead space minus predicted value, ΔSI_{III}: Slope of alveolar phase or phase III minus predicted value, IAH: Index of Alveolar Heterogeneity, IVE: Index of Ventilatory Efficiency). In every box, each curve corresponds to a different stage of the disease.

the clinical evolution of the disease. Unfortunately, there is no way to verify this hypothesis from our present data.

ROC curves analysis

In order to obtain an estimation of the diagnostic accuracy of each parameter, as well as to determine the cut-off for each stage level of the disease, we have performed a receiver operating characteristics (ROC) curve analysis. A

reliable definition of the disease and a good estimation of its prevalence, are the key elements for using ROC curves with a high degree of statistical power. Clear guidelines for staging COPD patients have been published by the GOLD group. The staging process is based on spirometry, a completely standardised and reproducible technique, which leaves only a narrow margin for ambiguity. Prevalence data of COPD stages have been recently published by Kornmann et al

(2003). Because their results are somewhat different from the internal prevalence calculated in our population, they have been used to obtain the cumulated prevalence for Positive and Negative Predicted Values and Likelihood Ratio calculations, except for Stage 0 level which does not represent a true prevalence of COPD disease, but simply a cut-off point for a reference population of ex-smokers (cessation for more than 10 years) with no clinical symptoms. Our population was not chosen to calculate the prevalence of COPD and that is why Stage Level 0 ROC Predicted Values and Likelihood Ratio do not correspond to the social prevalence of the disease. Furthermore, the stage-related variation of PPV and NPV may be influenced by their prevalence-dependence because we have used the cumulative prevalence in the ROC curve analysis. In contrast, sensitivity and specificity values, AUC and cut-off points, which are prevalence-independent are therefore valid for all stages, including stage 0.

The area under the ROC curve (AUC) is a global summary statistic of diagnostic accuracy. According to an arbitrary guideline (based on a suggestion by Swets in 1988), it is possible to distinguish between non-informative (AUC = 0.5), less accurate ($0.5 < 0.7 \leq 0.9$), highly accurate ($0.9 < AUC < 1$) and perfect tests (AUC = 1). It has been recognised that AUC is equivalent to the probability that a randomly drawn individual from the positive reference sample has a higher (or lower) test value than a randomly drawn individual from the negative reference sample. As AUC gives equal weighting to sensitivity and specificity, we have considered the point at which specificity equals sensitivity as the cut-off point for all stages. However, other criteria could be used if we consider, for instance, the misclassification cost (False Positive and False Negative Cost, CFP and CFN respectively), which require quantification of the consequences of misclassification (Greiner et al 2000). Without better information we have a priori considered that CFP = CFN, and we have given equal weight to sensitivity and specificity.

In a previous study Herholtz et al (2002) studied the diagnostic value of volumetric capnography parameters in a population of horses suffering from recurrent airways obstruction (a veterinary equivalent of COPD in horses). In their study the best accuracy was obtained for V_D^{Bohr}/V_T (AUC = 0.70, Se = Sp = 63%) in the staging between without symptoms and with symptoms horses, equivalent to our stage 0 for which we have obtained nearly identical results using $\Delta V_D^{Bohr}/V_T$ (AUC = 0.68, Se = Sp = 63%). However practically all our AUC have been over 0.7.

In conclusion, capnography parameters seem to be useful for staging COPD patients when spirometry is difficult to perform because of technical or clinical constraints. Capnography is an easy-to-perform technique that requires no special effort or particular knowledge of the manoeuvre, and can be performed in severely disabled patients. It can be used as an alternative test to diagnose the severity of functional impairment in these patients. The Index of Alveolar Heterogeneity seems to be the best parameter for use in clinical practice.

Disclosure

Author support

Abbreviations

ANOVA, Analysis of variance; AUC, Area Under the (ROC) Curve; BMI, Body Mass Index; CFN, False Negative Cost; CFP, False Positive Cost; COD, Coefficient of Determination; ΔSl_{III}, Slope of phase III real minus predicted; $\Delta V_D^{Bohr}/V_T$, V_D^{Bohr}/V_T ratio real minus predicted; F_{cov}, Fisher F for covariance analysis; $P_{ET}CO_2$, End tidal CO_2 partial pressure; FEV_1%pred, Forced Expiratory Volume in one second in percent of predicted; F_{group}, Fisher For inter-group comparison; IAH, Index of Alveolar Heterogeneity; IVE, Index of Ventilatory Efficiency; MMEF%pred, Maximal Mid Expiratory Flow in percent of predicted; NPV, Negative Predicted Value; P_{cov}, Probability associated to F_{cov}; P_{group}, Probability associated to F_{group} or χ^2_{group}; PIE, Pre-interface expirate; PLR, Positive Likelihood ratio; PPV, Positive Predicted Value; ROC, Receiver Operating Characteristics Curve; Se, Sensitivity; Sl_{III}, Slope of Phase III or alveolar phase in the capnogram; Sp, Specificity; VAE, Volume of alveolar ejection; $VCO_2(V)$, Curve of VCO_2 versus exhaled volume; V_D^{aw}, Airways Dead Space; V_D^{ser}, Serial Dead Space (includes instrumental dead space); V_ECO_2, CO_2 rejected by breath; V_T, Tidal Volume; χ^2_{group}, Chi squared from Kruskal-Wallis rank test;

References

Åström E, Niklason L, Drefeldt B, et al. 2000. Partitioning of dead space – a method and reference values in the awake human. *Eur Respir J*, 16:659–684.

ATS/ERS Standardization of Lung Function Testing: Standardization of Spirometry. http://www.thoracic.org/sections/publications/statements/pages/pfet/pft2.html

Blanch L, Lucangelo U, Lopez-Aguilar J, et al. 1999. Volumetric capnography in patients with acute lung injury: effects of positive end expiratory pressure. *Eur Respir J*, 13:1048–54.

Fletcher R, Jonson B. 1984. Deadspace and the single breath test for carbon dioxide during anaesthesia and artificial ventilation. *Br J Anaesth*, 6:109–76.

GOLD: Executive Summary: Global Strategy for the Diagnosis, Management, and Prevention of COPD. (Access 12 July 2006). URL: http://goldcopd.com/Guidelineitem.asp?l1=2&l2=1&intId=996

Greiner M, Gardner IA. 2000. Epidemiological issues in the validation of veterinary diagnostic tests. *Prev Vet Med*, 45:3–22.

Greiner M, Pfeiffer D, Smith RD. 2000. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Prev Vet Med*, 45:23–41.

Herholz C, Straub R, Lüthi S, et al. 2002. Validity of pulmonary function indices derived from the volumetric capnogram in horses with recurrent airway obstruction (RAO). *Res Vet Sci*, 72:141–46.

Hoffbrand BI. 1966. The expiratory capnogram: a measure of ventilation-perfusion inequalities. *Thorax*, 21:518–523.

Kars AH, Goorden G, Stijnen T, et al. 1995. Does phase 2 of the expiratory Pco2 versus volume curve have a diagnostic value in emphysema patients? *Eur Respir J*, 8:86–92.

Kars AH, Bogaard JM, Stijnen T, et al. 1997. Dead space and slope indices from the expiratory carbon dioxide tension-volume curve. *Eur Respir J*, 10:1829–36.

Koulouris NG, Latsi P, Imitroulis J, et al. 2001. Noninvasive measurement of mean alveolar carbon dioxide tension and Bohr's dead space during tidal breathing. *Eur Respir J*, 17:1167–74.

Kornmann O, Beeh KM, Beier J, et al. 2003. Newly diagnosed Chronic Obstructive Pulmonary Disease. *Respiration*, 70:67–75.

Roca J, Sanchis J, Agustí-Vidal A, et al. 1986. Spirometric reference values from a mediterranean population. *Bull Eur Physiopathol Respir*, 22:217–24.

Romero PV, Lucangelo U, Lopez Aguilar J, et al. 1997. Physiologically based indices of volumetric capnography in patients receiving mechanical ventilation. *Eur Respir J*, 10:1309–15.

Romero PV. 2001. Capnography: Basic Concepts. In: Gullo A, ed. Anesthesia, Pain, Intensive Care and Emergency Medicine. Vol 15. Milano: Springer-Verlag Italia; p.135–48.

Smidt U. 1976. Emphysema as possible explanation for the alteration of expiratory PO2 and PCO2 curves. *Bull Eur Physiopathol Respir*, 12:605–24.

Swets JA. 1988. Measuring the accuracy of diagnostic systems. Science 240:1285–93.

Toulou PP. 1966. Distribution of ventilation: clinical evaluation by rapid CO_2 analysis. *Dis Chest*, 49:139–45.

Wilschut FA, van der Grinten CPM, Lamers RJS, et al. 1999. Intrapulmonary gas mixing and the sloping alveolar plateau in COPD patients with macroscopic emphysema. *Eur Respir J*, 14:166–71.

Wolff G, Brunner JX. 1984. Series dead space volume assessed as the mean value of a distribution function. *Int J Clin Monit Comp*, 1:177–81.

Worth H. 1986. Expiratory partial pressure curves in the diagnosis of emphysema. *Bull Eur Physiopathol Respir*, 22:191–9.

You B, Mayeux D, Rkiek B, et al. 1992. La capnographie expiratoire dans l'asthme. *Rev Mal Resp*, 9:547–82.