

Reference values of pulmonary function tests for Canadian Caucasians

Carlos Gutierrez MD MSc¹, R Heberto Ghezso PhD², Raja T Abboud MD FRCPC³, Manuel G Cosio MD², John R Dill MD⁴, Richard R Martin MD FRCPC², Dan S McCarthy MD FRCPL⁵, John LC Morse MD FRCPC⁶, Noe Zamel MD FRCPC¹

C Gutierrez, RH Ghezso, RT Abboud, et al. Reference values of pulmonary function tests for Canadian Caucasians. *Can Respir J* 2004;11(6):414-424.

A multicentre, cross-sectional study was carried out in six centres across Canada to establish a national standard for pulmonary function tests using healthy, lifetime nonsmokers, with each centre aiming to test 10 men and 10 women from each decade from 20 to 80 years of age. Data from each centre were used to derive prediction equations for each centre, and pooled data from all centres (total: 327 women and 300 men) were used to derive Canadian predicted equations. The predictive models were compared with three widely used published models for selected tests. It was found that, in general, the equations modelled for each centre could be replaced by the models obtained when pooling all data (Canadian model). Comparisons with the published references showed good agreement and similar slopes for most tests. The results suggest that pulmonary function test results obtained from different centres in Canada were comparable and that standards currently used remain valid for Canadian Caucasians.

Key Words: *Body plethysmography; Pulmonary function tests; Reference values; Spirometry*

Pulmonary function tests (PFTs) are part of today's routine clinical assessment of respiratory function. As the number of centres, equipment and measuring methods grew, the need to standardize equipment, calibration and measurement techniques appeared necessary and guidelines were produced by the European Respiratory Society (1) and the American Thoracic Society (ATS) (2).

To interpret results, besides using properly calibrated equipment and standard techniques, reference values are required. Measured values of PFTs are usually compared with published standards obtained at sites other than the laboratories where the tests are performed. This can be justified for practical reasons, because it would be difficult, although probably ideal, to obtain local standards. We have produced a set of equations for PFTs that are valid across Canada by pooling values from several Canadian laboratories.

The purpose of the present study was to derive prediction equations for commonly used PFTs in the Canadian Caucasian

Les valeurs de référence des explorations fonctionnelles respiratoires pour les Canadiens de race blanche

Une étude multicentrique transversale a été effectuée dans six centres au Canada afin d'établir une norme nationale pour les explorations fonctionnelles respiratoires (EFR) à l'aide de personnes en santé qui n'avaient jamais fumé. Chaque centre visait à procéder à l'EFR chez 10 hommes et 10 femmes de chaque décennie entre 20 et 80 ans. Les données de chaque centre ont été utilisées afin de dériver des équations pronostiques pour chaque centre, et les données groupées de tous les centres (total de 327 femmes et de 300 hommes) ont été utilisées pour dériver les équations pronostiques canadiennes. Les modèles pronostiques ont été comparés à trois modèles publiés très utilisés pour des EFR sélectives. On a découvert qu'en général, les équations modelées pour chaque centre pouvaient être remplacées par les modèles obtenus après le regroupement de toutes les données (modèle canadien). Les comparaisons avec les références publiées ont démontré une bonne concordance et des pentes similaires pour la plupart des EFR. D'après les conclusions, les résultats des EFR obtenues dans les divers centres au Canada étaient comparables, et les normes actuellement utilisées demeurent valides pour les Canadiens de race blanche.

population, examine whether results of tests performed in different laboratories across the country were comparable and furthermore, test how our prediction equations, for selected tests, would perform when compared with published equations currently in use. We hypothesized that the standard formulas obtained using the Canadian population tested in pulmonary laboratories across the country using the techniques, equipment and calibration methods that the participating laboratories use in their day-to-day activities would be applicable throughout the country. For that reason, a research laboratory with specially trained personnel and/or equipment was not used; instead, we used the technicians who routinely perform lung function tests with the equipment available at each participating centre.

METHODS

Six centres across Canada selected a sample of subjects representative of a local healthy population. The participating centres were

¹Tri-Hospital Service (Mount Sinai – Toronto General – Women's College Hospitals), University of Toronto, Toronto, Ontario; ²Hôpital du Sacré-Coeur, University of Montreal, Montreal, Quebec; ³Vancouver General Hospital, University of British Columbia, Vancouver, British Columbia; ⁴Glance Bay Community Hospital, Glance Bay, Nova Scotia; ⁵Respiratory Hospital University of Manitoba, Winnipeg, Manitoba; ⁶St Joseph's Hospital, McMaster University, Hamilton, Ontario

Correspondence: Dr Noe Zamel, Mount Sinai Hospital, Room 474, 600 University Avenue, Toronto, Ontario M5G 1X5. Telephone 416-586-4473, fax 416-586-4507, e-mail noe@zamel.ca

selected because of their location and coverage. Subjects were selected from relatives and friends of patients and employees from the participating centres. They were all Caucasian, lifetime non-smokers, who were generally healthy and had no lung disease or respiratory symptoms. The centres were in Cape Breton, Nova Scotia; Montreal, Quebec; Toronto, Ontario; Hamilton, Ontario; Winnipeg, Manitoba; and Vancouver, British Columbia. Each centre was asked to test a quota of 10 nonsmoking, asymptomatic subjects per sex and age decade from 20 to 80 years of age. The subjects were all Caucasian with European parents. The subjects completed a simplified National Institute of Health-Division of Lung Diseases questionnaire, and if healthy, performed the following tests: forced expiratory and inspiratory maneuvers; lung volume determination either by body box or by helium dilution; single breath carbon monoxide (CO) diffusion test; airway resistance; and inspiratory and expiratory maximal mouth pressures, from residual volume (RV) and total lung capacity (TLC), respectively. The participants were interspersed with the regular laboratory clientele.

Equipment

The equipment used differed by centre and included a Morgan Model USA System for spirometry and lung diffusing capacity to CO (DL_{CO}) test (Morgan Scientific Inc, USA), and a Collins plethysmograph (Ferraris Respiratory, USA) for functional residual capacity (FRC); a spirometer model 211 and a body plethysmograph model 200 by CPI (Cardiopulmonary Instruments, USA); a diffusion analyzer model 88/CRT by Gould (Gould Medical Products Inc, USA) and pressure gauge (Marshall Town, USA); a Med Science wedge spirometer model 570 (Med-Science Electronics, USA), a body plethysmograph by PK Morgan (Morgan Scientific Inc) and a Med Science diffusion system, (Med-Science Electronics); a PK Morgan rolling seal spirometer with diffusing capacity module (Morgan Scientific Inc), and a Collins pressure plethysmograph (Ferraris Respiratory) for FRC; a Collins APE X DS/420 system (Ferraris Respiratory) and a Jaeger pressure plethysmograph (Metrolab Jaeger, Germany); and a Stead Wells spirometer (Braintree, USA) and a 1070 Medical Graphics System (MedGraphics, USA). All sites complied with ATS recommendations and the usual standardization and calibration procedures were followed (2-4). Ogilvie's technique was used for DL_{CO} measurement (5). Some centres did not perform certain tests because they were not part of their routine set of tests. Maximum inspiratory pressure and maximal expiratory pressure were measured using a transducer with a recorder. Two centres reported the maximal expiratory pressure at FRC and not at TLC.

Data analysis

A stratified analysis was done for each centre and a linear least square regression equation was modelled; these were called the Local equations. The variability of each PFT among centres was tested comparing the regression slopes by ANOVA. The data from all centres were pooled and regression models for each PFT, with age and height as predictive covariates, were obtained for each sex. Regression models for airway resistance and diffusion capacity based on the measured lung volume (body box and single breath helium dilution, respectively) instead of height were also determined, but they did not improve the regression models and were not analyzed in detail. Fitness of the models was determined by their coefficient of determination (r^2), standard error of the estimate (SEE or residual

SD = $S_{y|x}$) and analysis of residuals. If a factor did not significantly improve the predictive ability of the model, it was dropped. As a result, age or height for some tests were dropped from the final models. The final equations were called the Canadian equations. The data were explored extensively; transformations of the dependent variable into logarithms, as well as polynomials (square and cube), and square root transformations of the predictor variables were tested. Normality of the distributions of the dependent variables was tested by plotting cumulative residuals. A straight line indicated normally distributed data (6).

The local equations were tested for differences between slopes using ANOVA. The per cent predicted values of the Canadian and local equations were calculated and their differences were taken. No formal statistical analyses were performed for these values; a systematic deviation from 0 indicated a possible bias. To test how the Canadian equations would perform in classifying an observation as either normal or low compared with the local equations, the lower reference limits were calculated as the one-sided 5% lower confidence limit and the observations were classified accordingly. Measured values were considered low if they were less than that limit, otherwise they were considered normal. Lower limits of normality were calculated as the one-sided lower 5% confidence limit. The agreement beyond chance was then tested between such classified observations with the kappa statistic (7). Using Quanjer's (1,8) and Crapo's (9,10) formulas for both TLC and forced expiratory volume in 1 s (FEV_1), predicted values and their lower reference limits were calculated. The observations were then classified as either normal or low accordingly, and agreement beyond chance between Canadian and Crapo's predictions, and the Canadian and Quanjer's predictions for abnormality or normality was tested using the kappa statistic. The same procedure was followed for DL_{CO} with the Canadian, Crapo's (11) and Miller's prediction equations (12). Analysis was done using statistical software (SAS for Windows, version 8.0, SAS Institute, USA).

RESULTS

The total number of patients per 10 year age group ranged from 27 (men more than 70 years old) to 60. There were 327 women and 300 men. The numbers of female and male participants per centre are shown in Table 1. Some centres did not complete the quotas because they could not find any more subjects. Median body mass index (BMI) by centre among males ranged from 25 kg/m² to 26 kg/m², with the exception of one centre, where it was 23 kg/m² (centre 5). Median BMI for all men together was 25.2 kg/m², with first interquartile (25th percentile of the distribution) of 23.2 kg/m² and third interquartile (75th percentile of the distribution) of 27.3 kg/m². Over 50% of all men could be considered overweight. On the other hand, 5% had a BMI below 20.8 kg/m². Among women, the median BMI in different centres ranged from 22 kg/m² to 25 kg/m². For all women together, the median BMI was 23.6 kg/m², with first interquartile of 21.3 kg/m² and third interquartile of 26.2 kg/m². Approximately one-third of all women had a BMI greater than 25 kg/m², hence, they could be considered overweight. Ten per cent had a BMI of less than 20 kg/m².

Table 2 and Table 3 show the Canadian equations for women and men, respectively, with their respective r^2 , SEE and lower limit of normality value. For forced expiratory flow

TABLE 1
Number of participating subjects by decade and centre

Age	Centre 1 F:M	Centre 2 F:M	Centre 3 F:M	Centre 4 F:M	Centre 5 F:M	Centre 6 F:M	Total F:M
20 to <30	10:10	10:10	10:10	10:10	10:10	10:10	60:60
30 to <40	9:10	10:10	10:10	10:10	10:10	8:10	57:60
40 to <50	10:10	10:10	10:10	10:10	10:10	7:10	57:60
50 to <60	8:6	10:10	10:10	10:9	10:10	6:9	54:54
60 to <70	8:5	10:10	10:10	10:3	9:7	10:4	57:39
70 to <80	0:2	10:10	10:10	10:3	5:0	7:2	42:27
Total	45:43	60:60	60:60	60:45	54:47	48:45	327:300

F Female, M Male

TABLE 2
Canadian pulmonary function test prediction equations for adult females

	Intercept	Height	Age	S _{y x}	r ²	LLN
Total lung capacity (L)	-5.965	0.071	-0.007	0.583	0.47	
Vital capacity (L)	-3.597	0.050	-0.021	0.470	0.59	
Functional residual capacity (L)	-3.573	0.039		0.469	0.21	
Residual volume (L)	-2.314	0.020	0.015	0.378	0.26	
Forced vital capacity (L)	-3.335	0.049	-0.024	0.427	0.66	
FEV ₁ (L)	-1.901	0.037	-0.025	0.340	0.73	
FEF ₅₀ (L/s)	exp (1.778)		-0.010	0.286)	0.27	65%
FEF ₇₅ (L/s)	exp (1.232)		-0.021	0.351)	0.51	65%
FEV ₁ /forced vital capacity (%)	104.509	-0.089	-0.182	5.680	0.20	
Maximum inspiratory flow (L/s)	-0.622	0.047	-0.035	1.090	0.32	
MIF ₅₀ (L/s)	0.705	0.038	-0.037	1.117	0.31	
FEF ₅₀ /MIF ₅₀						
Lung diffusing capacity to CO (mL/min/mmHg)	-7.781	0.238	-0.153	3.469	0.49	
Lung diffusing capacity to CO (mL/min/mmHg)*	15.628	2.540*	-0.118	3.121	0.56	
CO diffusing coefficient	5.200		-0.014	0.697	0.11	
Airway resistance (cm H ₂ O/L/s)	4.636	-0.019		0.669	0.03	
Airway resistance (cm H ₂ O/L/s)†	2.016	-0.220†	0.005	0.663	0.04	
Specific airway conductance (L/s/cm H ₂ O/L)	-5.273	0.044		0.823	0.10	
Maximum inspiratory pressure (kPa)	95.578		-0.569	23.370	0.15	
MEP _{FRC} (kPa)	-8.121	0.649	-0.323	26.030	0.05	
MEP _{TLC} (kPa)	101.691		-0.657	21.960	0.21	

*This model was obtained using alveolar volume in place of height; †This model was obtained using thoracic gas volume in place of height. CO Carbon monoxide; FEF₅₀ Forced expiratory flow at 50% of expired vital capacity; FEF₇₅ forced expiratory flow at 75% of expired vital capacity; FEV₁ Forced expiratory volume in 1 s; LLN Lower limit of normality; MEP_{FRC} Maximum expiratory pressure measured at functional residual capacity; MEP_{TLC} Maximum expiratory pressure measured at total lung capacity; MIF₅₀ Maximum inspiratory flow at 50% of vital capacity; r² Coefficient of determination; S_{y|x} Standard error of the estimate

at 75% of expired vital capacity (FEF₇₅) and forced expiratory flow at 50% of expired vital capacity (FEF₅₀) transformation on the natural log scale yielded straight lines on the plot of cumulative residuals. To obtain the lower limit of normality for these two tests, the 5% lower limit (the one-sided 95% lower confidence limit of predicted values) was calculated. These values were transformed back to their original scale, divided by their predicted value and shown as a percentage.

The per cent of predicted values for each test was calculated using the local and Canadian equations by sex and centre, the

differences were obtained and their distributions were analyzed (data not shown). The differences between local and Canadian predicted values were less than five percentage points for most tests and a few were larger (eg, up to 30% for specific airway conductance). However, no obvious deviation in a specific direction (above or below 0) or consistently greater or smaller values was found for any centre. Hence, bias does not appear to be an explanation for those differences. And although some significant differences existed between centres (the magnitude of the difference depended on how well each model fit the data)

TABLE 3
Canadian pulmonary function test prediction equations for adult males

	Intercept	Height	Age	$S_{y x}$	r^2	LLN
Total lung capacity (L)	-8.618	0.090		0.817	0.37	
Vital capacity (L)	-5.897	0.069	-0.023	0.595	0.54	
Functional residual capacity (L)	-4.633	0.046		0.715	0.17	
Residual volume (L)	-2.443	0.020	0.021	0.470	0.33	
Forced vital capacity (L)	-5.473	0.067	-0.025	0.589	0.56	
FEV ₁ (L)	-2.832	0.047	-0.030	0.500	0.62	
FEF ₅₀ (L/s)	exp (2.036		-0.011	0.301)	0.26	61%
FEF ₇₅ (L/s)	exp (1.398		-0.020	0.361)	0.43	55%
FEV ₁ /Forced vital capacity (%)	109.396	-0.113	-0.210	5.821	0.24	
Maximum inspiratory flow (L/s)	-1.596	0.067	-0.047	1.619	0.27	
MIF ₅₀ (L/s)	0.145	0.055	-0.047	1.580	0.25	
FEF ₅₀ /MIF ₅₀	0.765		-0.002	0.268	0.01	
Lung diffusing capacity to CO (mL/min/mmHg)	-4.625	0.284	-0.246	5.368	0.43	
Lung diffusing capacity to CO (mL/min/mmHg)*	23.137	2.920*	-0.223	4.639	0.55	
CO diffusing coefficient	10.305	-0.023	-0.031	0.701	0.41	
Airway resistance (cm H ₂ O/L/s)	2.923	-0.010	0.004	0.553	0.03	
Airway resistance (cm H ₂ O/L/s)†	1.940	-0.230†	0.005	0.531	0.10	
Specific airway conductance (L/s/cm H ₂ O/L)	-10.837	0.079		1.769	0.08	
Maximum inspiratory pressure (kPa)	129.675		-0.653	28.190	0.12	
MEP _{FRC} (kPa)	151.409		-0.506	34.040	0.06	
MEP _{TLC} (kPa)	141.404		-0.636	30.280	0.07	

*This model was obtained using alveolar volume in place of height; †This model was obtained using thoracic gas volume in place of height. CO Carbon monoxide; FEF₅₀ Forced expiratory flow at 50% of expired vital capacity; FEF₇₅ forced expiratory flow at 75% of expired vital capacity; FEV₁ Forced expiratory volume in 1 s; LLN Lower limit of normality; MEP_{FRC} Maximum expiratory pressure measured at functional residual capacity; MEP_{TLC} Maximum expiratory pressure measured at total lung capacity; MIF₅₀ Maximum inspiratory flow at 50% of vital capacity; r^2 Coefficient of determination; $S_{y|x}$ Standard error of the estimate

the data were pooled and Canadian predictive models were obtained. As expected, the Canadian prediction equations had a better fit to all the data than did each centre's model. Although the mean predicted values may have differed significantly by centre, the slopes of the regression lines, ideally, should not. Differences between slopes (age and/or height) were tested. Indicator variables for the centres were included to test for significant differences between centres whenever the slopes were not different. However, these were not kept in the final models even if they improved the prediction ability of the regression, because keeping these would defeat the purpose of the project. For FEF₇₅, the age slopes were statistically significantly different for both sexes in one centre. For FEF₅₀, the age coefficient for men in that same centre was significantly different from the other centres. For forced expiratory volume as a percentage of forced vital capacity (FEV₁/FVC), the age coefficient was statistically significantly different for women; two centres, other than the one where age slopes for FEF₇₅ were different, showed significantly different slopes when compared with the other four centres. For DL_{CO}, age coefficients were significantly different for both women and men in the same centre where FEF₇₅ slopes were significantly different for both sexes (data not shown). None of the other sets of slopes showed statistically significant differences, which supported the use of the Canadian models to describe the data. This was supported further by the results of the tests for agreement.

Table 4 shows the classification of the observations as normal or low by the local and Canadian equations, kappa statistic, and positive and negative agreement statistics (P_{neg}) (13,14). For female subjects, negative agreement (results considered to be normal by both the Canadian and local predictive equations for all PFTs) was high, with P_{neg} greater than 0.95. Positive agreement (observations classified as low by both predictions) was high with the exception of RV and CO diffusing coefficient (K_{CO}). For RV, the Canadian equation ranked 5% (n=15) as low, whereas the local equation ranked 2% (n=7) as low, and only three subjects were ranked low by both equations. For K_{CO} , the Canadian equation ranked 3% (n=9) as low, whereas the local equation ranked 2% (n=5) low, and only one subject was ranked low by both equations. For male subjects, negative agreement was also excellent, with P_{neg} greater than 0.96. Positive agreement was also high, with the exception of maximum inspiratory flow (MIF) and K_{CO} . For MIF, the Canadian equation ranked 5% (n=12) as low and the local equation ranked 4% (n=10) low, but only three subjects were ranked low by both equations. For K_{CO} , the Canadian predictive equation ranked 4% (n=11) as low and the local predictive equation ranked 4% (n=10) as low, but only three subjects were considered low by both equations. With the above exceptions, the Canadian and local equations showed fair to excellent agreement in classifying the observations as either normal or low.

TABLE 4
Agreement in classification of female and male subjects as either low or normal according to the local and Canadian equations

Female	a	b	c	d	Total	Kappa	P _{pos}	P _{neg}
Total lung capacity	7	6	6	307	326	0.52	0.54	0.98
Vital capacity	7	8	9	303	327	0.42	0.45	0.97
Functional residual capacity	9	3	4	310	326	0.71	0.72	0.99
Residual volume	3	4	12	307	326	0.25	0.27	0.97
Forced vital capacity	7	5	8	307	327	0.50	0.52	0.98
FEV ₁	6	4	8	309	327	0.48	0.50	0.98
FEF ₅₀	8	5	5	268	286	0.60	0.62	0.98
FEF ₇₅	7	1	7	269	284	0.62	0.64	0.99
FEV ₁ /forced vital capacity	14	4	10	298	326	0.64	0.67	0.98
MIF	6	3	3	280	292	0.65	0.67	0.99
MIF ₅₀	5	3	2	207	217	0.65	0.67	0.99
FEF ₅₀ /MIF ₅₀ *								
Lung diffusing capacity to CO	5	6	5	308	324	0.46	0.48	0.98
CO diffusing coefficient	1	4	8	269	282	0.13	0.14	0.98
Airway resistance	2	0	1	219	222	-0.01	0.80	1.00
Specific airway conductance	3	3	1	215	222	0.59	0.60	0.99
Maximum inspiratory pressure	5	1	0	217	223	0.91	0.91	1.00
MEP _{FRC}	2	1	1	114	118	0.66	0.67	0.99
MEP _{TLC}	3	0	2	101	106	0.74	0.75	0.99
Male	a	b	c	d	Total	Kappa	P _{pos}	P _{neg}
Total lung capacity	8	1	4	287	300	0.79	0.80	0.99
Vital capacity	10	2	6	282	300	0.70	0.71	0.99
Functional residual capacity	6	2	1	291	300	0.79	0.80	0.99
Residual volume	9	4	6	281	300	0.63	0.64	0.98
Forced vital capacity	10	0	5	285	300	0.79	0.80	0.99
FEV ₁	6	2	10	282	300	0.48	0.50	0.98
FEF ₅₀	9	1	6	252	268	0.71	0.72	0.99
FEF ₇₅	5	4	6	252	267	0.48	0.50	0.98
FEV ₁ /Forced vital capacity	8	4	7	281	300	0.57	0.59	0.98
MIF	3	7	9	218	237	0.24	0.27	0.96
MIF ₅₀	5	3	2	207	217	0.65	0.67	0.99
FEF ₅₀ /MIF ₅₀ *								
Lung diffusing capacity to CO	10	4	3	280	297	0.73	0.74	0.99
CO diffusing coefficient	3	7	8	248	266	0.26	0.29	0.97
Airway resistance	1	1	1	205	208	0.50	0.50	1.00
Specific airway conductance	0	0	0	208	208	*		
Maximum inspiratory pressure	6	1	2	197	206	0.79	0.80	0.99
MEP _{FRC}	2	0	1	115	118	0.80	0.80	1.00
MEP _{TLC}	4	1	1	80	86	0.79	0.80	0.99

*No equation. a Low by both Canadian and local equations; b Low by local but not by Canadian equation; c Low by Canadian but not by local equations; d Normal by both Canadian and local equations; CO Carbon monoxide; FEF₅₀ Forced expiratory flow at 50% of expired vital capacity; FEF₇₅ forced expiratory flow at 75% of expired vital capacity; FEV₁ Forced expiratory volume in 1 s; MEP_{FRC} Maximum expiratory pressure measured at functional residual capacity; MEP_{TLC} Maximum expiratory pressure measured at total lung capacity; MIF Maximum inspiratory flow; MIF₅₀ Maximum inspiratory flow at 50% of vital capacity; P_{neg}=2d/(N-a+d), negative agreement statistic; P_{pos}=2a/(a+b+a+c), positive agreement statistic

Agreement between Canadian and selected published prediction equations for FEV₁, TLC and DL_{CO}

Table 5 shows the number of observations for each cell of the agreement table, total number of subjects, kappa, and positive agreement statistic and P_{neg} by sex. With Quanjer's prediction equations, negative agreement for all was excellent.

For FEV₁ in female subjects, the Canadian and Quanjer's equations showed poor positive agreement; only one subject (0.3%) was ranked low by Quanjer's equation only, whereas 14 subjects (4%) were ranked low by the Canadian equation. For male subjects, positive agreement was poor too, with only one of the subjects having been ranked low by Quanjer's

TABLE 5
Agreement in classification of subjects as either low or normal by the Canadian equations compared with Quanjer's, Crapo's and Miller's equations for selected pulmonary function tests

Model	Test	Sex	a	b	c	d	Total	Kappa	P _{pos}	P _{neg}
Quanjer's	FEV ₁	F	1	0	13	313	327	0.13	0.13	0.98
	FEV ₁	M	1	0	15	284	300	0.11	0.12	0.97
	Total lung capacity	F	4	0	9	313	326	0.46	0.47	0.99
	Total lung capacity	M	11	0	1	288	300	0.96	0.96	1.00
Crapo's	FEV ₁	F	5	0	9	313	327	0.52	0.53	0.99
	FEV ₁	M	10	0	6	284	300	0.76	0.77	0.99
	Total lung capacity	F	8	4	5	309	326	0.63	0.64	0.99
	Total lung capacity	M	3	0	9	288	300	0.39	0.40	0.98
	Lung diffusing capacity to CO	F	10	43	0	271	324	0.30	0.32	0.93
	Lung diffusing capacity to CO	M	13	33	0	251	297	0.45	0.44	0.94
Miller's	Lung diffusing capacity to CO	F	4	0	6	314	324	0.56	0.57	0.99
	Lung diffusing capacity to CO	M	8	0	5	284	297	0.75	0.76	0.99

a Low by both Canadian and reference equation; b Low by reference but normal by Canadian equation; c Low by Canadian but normal by reference equation; d Normal by both Canadian and reference equation; CO Carbon monoxide; F Female; FEV₁ Forced expiratory volume in 1 s; M Male; P_{neg}=2d/(N-a+d), negative agreement statistic; P_{pos}=2a/(a+b+a+c), positive agreement statistic

equation, 15 subjects (5%) ranked low by the Canadian equation, one subject was ranked low by both. For TLC in female subjects, positive agreement was fair to good. The Canadian equation ranked 13 subjects (4%) low and Quanjer's equation indicated four subjects (2%) were low; these four were also ranked low by the Canadian equation. For male subjects, positive agreement was also excellent; the Canadian and Quanjer's equations both ranked 11 subjects (4%) low, while Quanjer's ranked an additional subject low, and the remaining subjects were ranked normal by both equations.

With Crapo's prediction equations for FEV₁ and TLC, negative agreement for all was excellent. For FEV₁ in female subjects, positive agreement was fair with 14 (4%) of the subjects ranking low by the Canadian equation, whereas five (2%) subjects were ranked low by Crapo's equation. For male subjects, positive agreement for FEV₁ was excellent with 16 (5%) of the subjects ranked low by the Canadian equation and 10 (3%) subjects ranked low by Crapo's equation (all 10 subjects were ranked low by the Canadian equation as well). For TLC in female subjects, positive agreement was fair to good: 12 (4%) of the subjects were low by Crapo's equation and 13 (4%) subjects were ranked low by the Canadian equation, with eight subjects ranked low by both equations. For male subjects, positive agreement for TLC was poor: three (1%) of the subjects were ranked low by Crapo's equation, whereas 12 (4%) subjects were ranked low by the Canadian equation, including the three observations considered low by Crapo's equation.

A good agreement indicates that the present equations are as good as the published equations currently in use for the purpose of classifying a new subject as either normal or low. As expected, the Canadian equations ranked approximately 5% of the subjects low. If Quanjer's equations were applied to this population, poor agreement for FEV₁ among both female and male subjects would be obtained and less than 1% of all observations would have been considered low. For TLC, the results would be close to those obtained with the Canadian equation. If Crapo's equations were used, approximately one-half the number females classified low by the Canadian equation would be misclassified as normal for FEV₁; a similar situation would occur for TLC.

With Crapo's FEV₁ model in male subjects, the results would be similar to those of the Canadian equation, whereas for TLC, Crapo's equation would misclassify approximately two-thirds of those ranked low by the Canadian equation as normal.

For DL_{CO}, poor agreement (low positive agreement) was found for both female and male subjects with Crapo's prediction equations (Table 5). An additional 10% to 15% of subjects would be considered low by Crapo's equation compared with the Canadian equations. Predicted values by Crapo's equation were greater than the Canadian predicted values by a mean of 10% for female and 8% for male subjects. When the Canadian equations were compared with Miller's equations, results with the two predictions were similar. For female subjects, positive agreement was fair to good: four (1%) subjects were ranked low by Miller's equation and 10 (3%) subjects were ranked low by the Canadian equation (four subjects were ranked low by both equations). For male subjects, positive agreement was excellent: eight (3%) subjects were ranked low by Miller's equation, 13 (4%) subjects were considered low by the Canadian equation, and eight subjects were classified low by both equations. Although a large variation in DL_{CO} in different centres may be expected (15), the results with the Canadian prediction values agreed quite well with Miller's equation, but not with Crapo's. It should be noted that Miller et al (12) compared various prediction equations available at that time with their own and found that Crapo's predictions gave higher results than all other equations. Crapo's higher values could be related to possible ethnic differences (Crapo's subjects were Mormons or Seventh-day Adventists), or possibly, to the effect of residence at altitude in Salt Lake City, Utah (9).

Figures 1 and 2 show regression lines using the Canadian, Quanjer's, and Crapo's prediction models for FEV₁ and TLC. It can be seen that, in general, for FEV₁ and TLC, the results with the Canadian models were somewhat greater than with Quanjer's and Crapo's. However, for FEV₁ the three regression lines have very similar slopes but with different intercepts (Figure 1); values with the Canadian models were slightly higher than those with the other two models, and were closer to Crapo's than Quanjer's predicted values. The Canadian

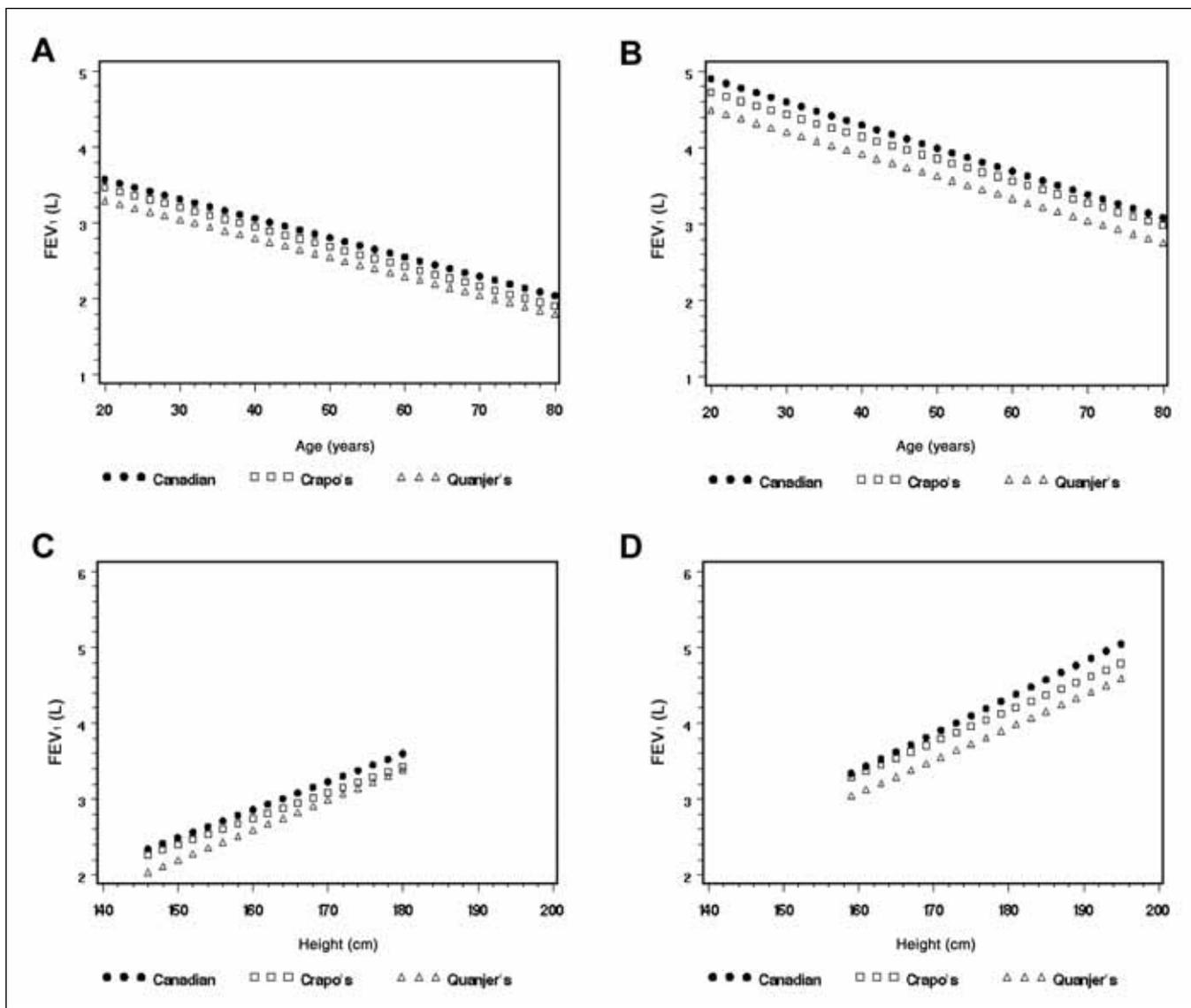


Figure 1 Predicted forced expiratory volume in 1 s (FEV_1) by height, age and sex using the Canadian, Crapo's and Quanjer's equations. (A) FEV_1 plotted against age for 20- to 80-year-old women 165 cm in height. (B) FEV_1 plotted against age for 20- to 80-year-old men 175 cm in height. (C) FEV_1 plotted against height for 45-year-old women 146 cm to 180 cm in height. (D) FEV_1 plotted against height for 45-year-old men 158 cm to 196 cm in height. In general, the slopes were similar with the three models; however, the predicted values with the Canadian equation, although slightly greater, were closer to Crapo's than to Quanjer's predicted values

predicted values for TLC decreased slightly with age and were slightly higher than those with Quanjer's and Crapo's, which were very similar (Figure 2). When comparing predicted DL_{CO} values with the Canadian, Crapo's, and Miller's models, the slopes plotted by age (panels A and B) were similar but the intercepts were higher for Crapo's equation than for the Canadian and Miller's equations (Figure 3, panels A and B). When predicted values were plotted by height for 45-year-old subjects (Figure 3, panels C and D), the slopes for female subjects were similar; however, for male subjects, there was a steeper increase in DL_{CO} with increasing height with Crapo's model than with the other two regression models; the slope was the lowest with Miller's equation. Again, predicted values with Crapo's models were greater than with the Canadian and Miller's, and those with Miller's models were closer but slightly lower than those obtained with the Canadian models.

DISCUSSION

We have modelled equations for several PFTs based on data in healthy, Caucasian lifetime nonsmokers from six centres across Canada. An advantage to having national standards is having models valid across the country that would better describe the Canadian population. We found that, although there were slight differences among some centres, our prediction equations appear valid across the country. Our sample was relatively large and although the participants were not chosen at random, we believe it to be representative of the target population, as there was no obvious bias in the selection process. Our subjects were assumed to be healthy based on their responses to the questionnaire. Nonetheless, a relatively large proportion of people in our sample was overweight and a very small proportion was underweight. Indeed, 25% of men and 22% of women had BMIs greater than 27 kg/m^2 and there were some with BMIs

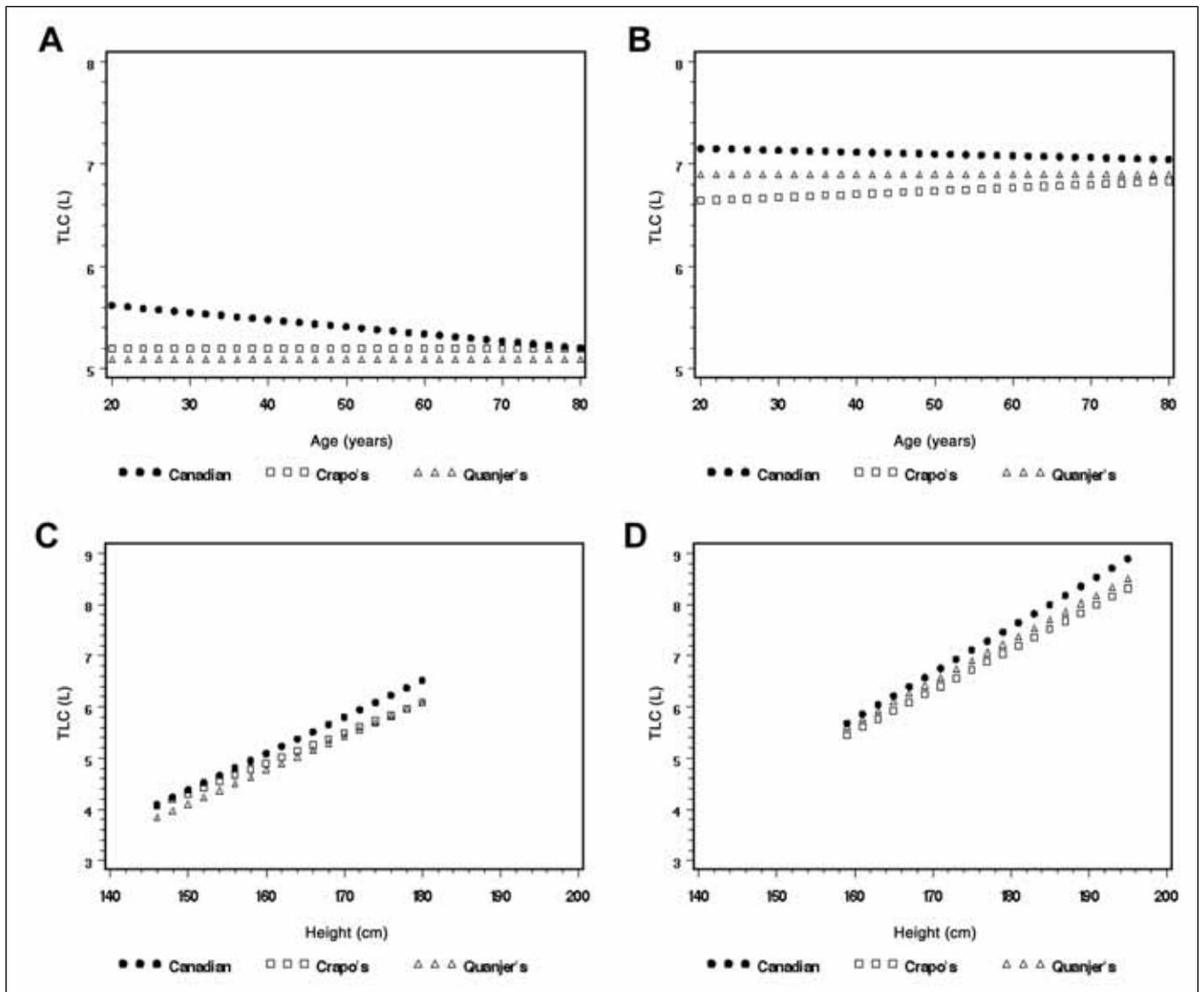


Figure 2 Predicted total lung capacity (TLC) by sex, age and height using the Canadian, Crapo's and Quanjer's equations. (A) TLC plotted against age for 20- to 80-year-old women 165 cm in height. (B) TLC plotted against age for 20- to 80-year-old men 175 cm in height. (C) TLC plotted against height for 45-year-old women 146 cm to 180 cm in height (D) TLC against height for 45-year-old men 158 cm to 196 cm in height. When plotted by age, the Canadian model indicated that TLC decreased as age increased and the predicted values were greater than with Quanjer's or Crapo's models. TLC for women remained unchanged with age for both Quanjer's and Crapo's predictions (age was not a factor in these models). For men, TLC increased with increasing age according to Crapo's model, whereas according to Quanjer's model it remained unchanged. However, for TLC plotted against height, all three regressions had similar slopes with the Canadian slope being greatest than the other two, which were similar

greater than 30 kg/m^2 . Although lung function may decrease at both extremes of weight (16), BMI was not predictive in any of the tested models.

The ATS set norms for the performance and reporting of the following tests: FVC, FEV_1 , DL_{CO} , forced expiratory flow at 50% to 75% of expired vital capacity, FEF_{50} , FEF_{75} , TLC, RV and FRC (2); and later updated those for vital capacity, FVC, FEV_1 , forced expiratory flow at 50% to 75% of expired vital capacity, flow and flow-volume curves (3), and DL_{CO} (4). Although not standardized by ATS, some of the tests in the present study used similar equipment and were obtained using the same maneuvers as the standardized tests, thus, their reproducibility should approach those of standardized

tests. However, some of the tests studied, such as airways resistance or maximal mouth pressures, are not well standardized, difficult to perform by the patient and subject to variations in reading by the technician or even computers. Hence, greater variability for these tests among institutions is to be expected.

PFTs are used in the clinical setting to support a diagnosis of an intrinsic respiratory, muscular or neurological condition affecting the respiratory system, and to evaluate the degree of functional impairment. Usually, a set of PFTs is obtained as part of the initial diagnostic evaluation, and will also serve as a baseline for comparison with subsequent tests to assess the evolution of the condition and, more frequently, to evaluate the response to therapy. PFTs are also obtained to detect possible

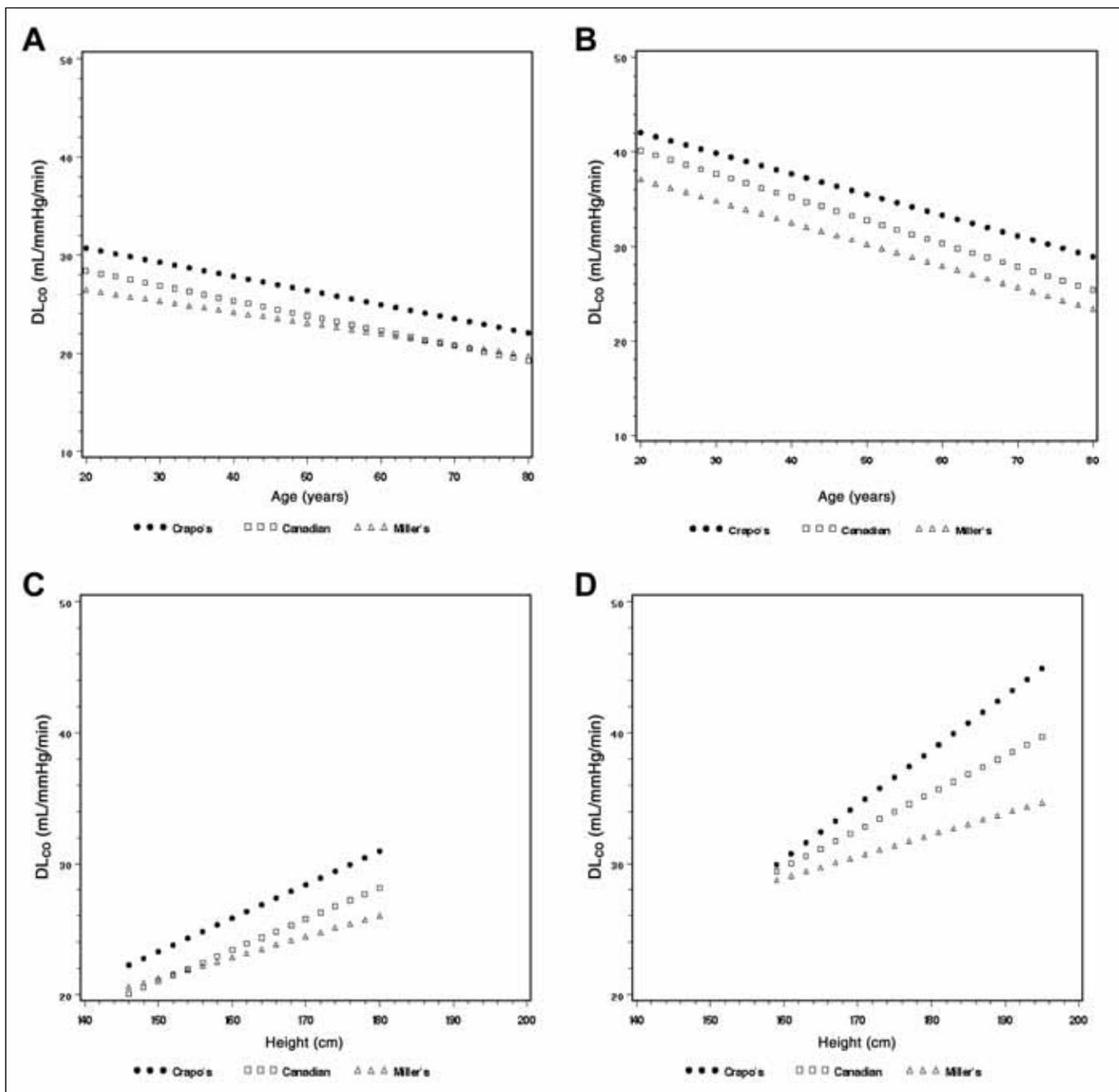


Figure 3) Predicted lung diffusing capacity to carbon monoxide (DL_{CO}) by sex, age and height using the Canadian, Crapo's, and Miller's equations. (A) DL_{CO} plotted against age for women 165 cm in height and aged 20 to 80 years. (B) DL_{CO} plotted against age for men of 175 cm in height aged 20 to 80 years. (C) DL_{CO} plotted against height for 45-year-old women 146 cm to 180 cm in height. (D) DL_{CO} for 45-year-old men 158 cm to 196 cm in height. When plotted by age, predicted DL_{CO} values decreased with increasing age. Although the slopes did not appear very different, the values tended to converge between the Canadian and Miller's models, and diverge from Crapo's model as age increased. In women, DL_{CO} increased with increasing height with slopes similar to Crapo's slopes (C). For men, the predicted values plotted against height were quite different with each model. For the lower heights, the predicted values were close, but as height increased so did DL_{CO} ; however, it increased at a faster rate with the Canadian model than with Miller's model, and much faster with Crapo's model (D). In general, compared with the Canadian model, Crapo's model yielded higher DL_{CO} values, while Miller's model yielded lower DL_{CO} values

respiratory side effects of treatments for nonrespiratory conditions. When used in support of a diagnosis, prediction equations are of value in classifying the result as normal or abnormal with respect to a reference population. Hence, although PFTs are measured on a continuous scale, they are used as bivariate data, ie, normal versus low. To classify data

this way, we used the calculated lower limits of normality. We tested for agreement in such a classification between the local and Canadian models using the kappa statistic. The results showed that the Canadian predictions were as good as the local equations for classifying a result in that manner. Therefore, we think it is valid to use the Canadian models.

We found differences in slopes for FEV_1/FVC , FEF_{75} and DL_{CO} which would not favour comparing results obtained in one centre with those obtained in a second centre. The differences in the intercepts may be explained by differences among subjects, equipment, technician, other nonaccounted factors and/or a mixture or interaction of these factors. The differences in slopes are more difficult to explain because slope coefficients quantify the linear relationship between the predictor and dependent variables, and because there is no existing obvious biological reason for a change in the extent of that association, we think that measurement-related factors were more likely involved. But we have to be cautious and consider the predicting ability of each model. A small r^2 with a large SEE would imply that a model explains or predicts very little of the dependent variable, as was the case for the FEV_1/FVC ratio. Therefore, the reason for the differences in slopes of that ratio may lie elsewhere. For FEF_{75} , the linear model fit the data much better; therefore, a good part of the differences is more likely related to measurement. Among female subjects, four centres had age slopes for $\log FEF_{75}$ that were not significantly different. There were two centres with much steeper slopes than all other centres. For male subjects only, one of the centres, had a much steeper slope with age (it was also different for the females). The other five centre slopes were similar. We have not found an explanation for these differences.

We found that if the testing centre was included in the equation coded as an indicator (dummy) variable, we could improve the predictive ability of some models (data not shown). However, we considered that the nature of the study – collecting data from laboratories spread across the country using standard methods performed with the equipment and personnel that carry out the routine day-to-day testing – would increase the variability and because our objective was to model prediction formulas that could be used in any laboratory in Canada that uses standard equipment and follows the same guidelines, we accepted this drawback. We considered it more important for no bias to be apparent.

No linear equation could be modelled for $FEF_{50}:MIF_{50}$ for female subjects. We reported the results for this ratio for male subjects. However, one must be cautious using this latter model because its predictive ability is very poor.

Linear models have limitations. The prediction of values near the extremes of the regression line are less accurate. Hence, exponential models have been introduced claiming better fit to the data. In general, this is true for PFTs models where adolescents are included in the sample. The explanation is that adolescents' lungs are still growing and this growth peaks between 18 and 25 years (17,18). For adult populations, there appears to be no advantage in using exponential models (16). Ghezzi (19) did a detailed analysis of several nonlinear formulas of PFTs. He did not find major differences for elderly subjects, but noticed a marked improvement in the 16- to 25-year-old group. In general, we use different formulas for adolescents than for adults with a cut-off at roughly 20 to 25 years of age. Models for instant flows, such as FEF_{50} and FEF_{75} , were the only PFTs among our data that showed an improvement compared with the simple linear models.

The results based on the Canadian predictions did not show perfect agreement when compared with Quanjer's and Crapo's

models for TLC and FEV_1 . Predicted values for FEV_1 and TLC were slightly higher with the Canadian equation. For DL_{CO} , Crapo's predicted values were higher by approximately 10% for male and 8% for female subjects than the Canadian predicted values, and they were also higher than Miller's predictions. However, slopes for DL_{CO} among female subjects, and for TLC and FEV_1 , for both female and male subjects, were rather similar, but the intercepts were different. That is, there were factors that shifted the intercept for the different models, but in general, the magnitude of the association between age and height with the dependent variables was the same in the tested models. Predicted DL_{CO} values among male subjects appeared different; the values increased with height at a faster rate with Crapo's equation than with the Canadian model, whereas with Miller's equation they increased at a slower rate than with the Canadian model.

No detailed analysis was done of the models predicting DL_{CO} and airway resistance based on lung volume as a predictive variable. Using r^2 , we found that the models for DL_{CO} based on single breath TLC did not appear better than those based on height and age. This may be true for normal subjects, but probably not for subjects with previous lung resections, for example; however, this would have to be tested. Linear regression models for $FEF_{50}:MIF_{50}$ ratio and airway resistance with height and age or volume have very little predictive ability.

In general, the equations more widely used in Canada were generated in the United States in Caucasian populations. Although there are no reasons to expect the Caucasian population in Canada to be different from that in the United States, we were interested in obtaining prediction equations from a Canadian population, because equations derived in this manner should more closely reflect the target population. We have modelled predictive equations for Canadian Caucasians that are valid across the country. When we tested our models for selected PFTs with Crapo's, Miller's and Quanjer's prediction equations, we found them to be fairly similar overall, although there were some differences. The similarities are not surprising because Caucasians in North America of European origin would be expected to be similar to each other and also similar to European Caucasians.

REFERENCES

1. Standardized lung function testing. Bull Eur Physiopathol Respir 1983;19(Suppl 5):1-95.
2. Ferris BG. Epidemiology Standardization Project (American Thoracic Society). Am Rev Respir Dis 1978;118:1-120.
3. American Thoracic Society. Standardization of spirometry – 1987 update. Am Rev Respir Dis 1987;136:1285-98.
4. American Thoracic Society. Single-breath carbon monoxide diffusing capacity (transfer factor). Recommendations for a standard technique. Am Rev Respir Dis 1987;136:1299-307.
5. Blakemore WS, Forster RE, Morton JW, Ogilvie CM. A standardized breath holding technique for the clinical measurement of the diffusion capacity of the lung for carbon monoxide. J Clin Invest 1957;36:1-17.
6. Glanz SA, Linker BK, eds. Primer of Applied Regression and Analysis of Variance. Toronto: McGraw-Hill Inc, 1990:125-30.
7. Tammemagi MC, Frank JW, Leblanc M, Artsob H, Streiner DL. Methodological issues in assessing reproducibility – a comparative study of various indices of reproducibility applied to repeat ELISA serologic tests for Lyme disease. J Clin Epidemiol 1995;48:1123-32.

8. Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J Suppl* 1993;16:5-40.
 9. Crapo RO, Morris AH, Clayton PD, Nixon CR. Lung volumes in healthy nonsmoking adults. *Bull Europ Physiopath Resp* 1982;18:419-25.
 10. Crapo RO, Morris AH, Gardner RN. Reference spirometric values using techniques and equipment that meet ATS recommendations. *Am Rev Respir Dis* 1981;123:659-64.
 11. Crapo RO, Morris AH. Standardized single breath normal values for carbon monoxide diffusing capacity. *Am Rev Respir Dis* 1981;123:185-9.
 12. Miller A, Thornton JC, Warsaw R, Anderson H, Teirstein AS, Selikoff IJ. Single breath diffusing capacity in a representative sample of the population of Michigan, a large industrial state. Predicted values, lower limits of normal, and frequencies of abnormality by smoking history. *Am Rev Respir Dis* 1983;127:270-7.
 13. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990;43:543-9.
 14. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990;43:551-8.
 15. Crapo RO, Jensen RL, Wanger JS. Single-breath carbon monoxide diffusing capacity. *Clin Chest Med* 2001;22:637-49.
 16. Dockery DW, Ware JH, Ferris BG Jr, et al. Distribution of forced expiratory volume in one second and forced vital capacity in healthy, white, adult never-smokers in six U.S. cities. *Am Rev Respir Dis* 1985;131:511-20.
 17. Schoenberg JB, Beck GJ, Bouhuys A. Growth and decay of pulmonary function in healthy blacks and whites. *Respir Physiol* 1978;33:367-93.
 18. Knudson RJ, Lebowitz MD, Holberg CJ, Burrows B. Changes in the normal maximal expiratory flow-volume curve with growth and aging. *Am Rev Respir Dis* 1983;127:725-34.
 19. Ghezzi RH. Normalization of lung function tests for epidemiological studies. PhD thesis. Department of Epidemiology and Biostatistics, McGill University, Montreal, Quebec, 1982.
-
-



Hindawi
Submit your manuscripts at
<http://www.hindawi.com>

