



Effect of changing reference equations for spirometry interpretation in Thai people

Warawut Chaiwong¹, Sureeporn Uthaikhup², Chalerm Liwsrisakun¹, Chaicharn Pothirat¹

¹Division of Pulmonary, Critical Care and Allergy, Department of Internal Medicine, Faculty of Medicine, ²Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand

Contributions: (I) Conception and design: P Therapist; (II) Administrative support: All authors; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: P Therapist; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Prof. Chaicharn Pothirat, MD, FCCP. Division of Pulmonary, Critical Care and Allergy, Department of Internal Medicine, Faculty of Medicine, Chiang Mai University, 110 Inthavaroros Rd. Sriphum, Maung Chiang Mai, Chiang Mai 50200, Thailand. Email: chaicharn.p@cmu.ac.th.

Background: Use of different spirometry reference equations can lead to misinterpretation of spirometry results. Currently, no study has been carried out to determine if there is any discordance between the interpretation of spirometry data using the Global Lungs Initiative (GLI) 2012, the Third National Health and Nutrition Examination Survey (NHANES III), the Knudson, and the Siriraj reference equations in Thai people. Therefore, this study aimed to determine differences and agreements regarding spirometry results using the GLI2012, NHANES III, Knudson, and Siriraj reference equations in a Thai population.

Methods: A retrospective study was carried out using 2,492 spirometry results from Thai people (age ranged 8–80 years old) who were referred by their physicians for spirometry testing at the Lung Health Center, Maharaj Nakorn Chiang Mai Hospital, Chiang Mai, Thailand between January 2005 and December 2015. The spirometry results were interpreted using four reference equations. Respiratory function patterns were classified as normal, spirometric restriction, obstruction, or mixed defect. The severity of abnormal patterns were classified as mild, moderate, moderately severe, severe or very severe. A Nominal Analysis of Variance (NANOVA) test was used to analyze differences in diagnosis and classification for the degree of severity. The Kappa statistic was used to determine agreements of spirometry results between four reference equations.

Results: The differentiation in both spirometry diagnosis and the severity classification were significant across the four reference equations ($P < 0.001$). However, the levels of agreement among the four reference equations were moderate to good (Kappa values ranged from 0.56 to 0.77 for the diagnosis and 0.67 to 0.82 for the severity).

Conclusions: The results of this study suggest that the use of the GLI2012, NHANES III, Knudson, and Siriraj reference equations alters the interpretation of spirometry data in this population of Thai adults.

Keywords: Spirometry; equation; pulmonary function; agreement; respiratory

Submitted Jul 08, 2018. Accepted for publication Sep 06, 2018.

doi: 10.21037/jtd.2018.11.06

View this article at: <http://dx.doi.org/10.21037/jtd.2018.11.06>

Introduction

Spirometry is a medical screening tool that measures various aspects of pulmonary function (1). It is the most common lung function test and abnormal spirometry results play an important role in both the diagnosis

of respiratory diseases and the classification of their severity (2). However, reliable interpretation of spirometry results depends on spirometry reference equations used (3,4). Therefore, use of an inappropriate reference can lead to errors in both the diagnosis of respiratory diseases

and the classification of their severity from abnormal spirometry results (5,6).

The Third National Health and Nutrition Examination Survey (NHANES III) reference has been widely used in many clinical laboratories and clinical trials (2). However, these references provide the most appropriate equation for the Caucasian population (7). Another reference equation has been published by Knudson *et al.* (8). The Knudson reference equation provides the equation most appropriate to a non-Mexican-American population and has also been widely used over several years (2,8,9). In 2000 the Siriraj reference equations were established for the Thai population (10). However, the NHANES III and Knudson reference values are also fairly commonly used in Thai situations (11,12). A tentative explanation may be that underestimated ventilatory defects are often clinically observed in Thai people when interpreted with Siriraj reference equations when compared to the NHANES III and Knudson reference equations. More recently, the Global Lung Initiative (GLI) 2012 reference equations have been established by the European Respiratory Society (ERS) and are considered to be the first multi-ethnic all age reference equations for spirometry (13). The GLI2012 equation has been recommended as being the most appropriate of the spirometric predicted equations for every country including Thailand.

Evidence suggested that the variation between ethnic groups is an important issue for the interpretation of spirometry (4). Previous studies showed that spirometric prediction equations proposed for Asian populations were found to be significantly different from those equations recommended for use with Caucasians (14,15). Additionally, recent studies have showed that the spirometry predicted values by GLI2012 reference equations were significantly different from those suggested by the other Caucasian reference equations (16,17). Currently the GLI2012 equations have been recommended as global reference equations for spirometry and for every country. However, the effect of applying the GLI2012 equations to Thai people remains unknown. In addition, interpretation of spirometry using the commonly used reference equations (NHANES III, Knudson, Siriraj) and the GLI2012 may be similar or different from one another. Further investigation of use of these four reference equations in Thai people is still required. Therefore, this study aimed to investigate any discordance in interpretation of spirometry data using the GLI2012, NHANES III, Knudson, and Siriraj reference equations for Thai people.

Methods

A retrospective study of spirometry results was carried out using data from Thai adults (age 18–80 years old) who were referred by their physicians for spirometry testing at the Lung Health Center, Maharaj Nakorn Chiang Mai Hospital, Chiang Mai, Thailand between January 2005 and December 2015. There was limited data available for childhood subjects aged <18 years old at the Lung Health Center. Pre-bronchodilator spirometry results from the first visit of each subject were included for this study. The spirometry was performed using a standard spirometer (Vmax series 22, Sensormedics, USA) and all spirometry results were measured according to the American Thoracic Society (ATS)/ERS 2005 guidelines (18). The study was approved by the Research Ethics Committee of the Faculty of Medicine, Chiang Mai University [Institutional Review Board (IRB), approval number: Med-2559-04324].

Demographic data including age, gender, height, and weight were collected. Spirometry parameters were also recorded [forced vital capacity (FVC), forced expiratory volume in the first second (FEV₁), and percentage of the ratio of FEV₁ to FVC (%FEV₁/FVC)]. The predicted and lower limit of normal (LLN) values of FVC, FEV₁, and %FEV₁/FVC were calculated from the GLI2012 (Southeast Asian sub-group), NHANES III, Knudson, and Siriraj reference equations. For Asian population, the predicted and LLN values for FVC and FEV₁ estimated using the NHANES III and Knudson reference equations were multiplied by a conversion factor of 0.88 and 0.94, respectively (1). No conversion factor was required for the Siriraj and GLI2012 equations because they were derived from Thai populations and Southeast Asian sub-group populations, respectively.

The spirometry results collected were classified into four patterns: normal, obstructive, spirometric restrictive, and mixed restrictive and obstructive. The criteria for each pattern were defined as follows: %FEV₁/FVC value > LLN and FVC > LLN for a normal pattern, %FEV₁/FVC < LLN but the FVC ≥ LLN for an obstructive defect, %FEV₁/FVC was ≥ LLN but the FVC < LLN for a spirometric restrictive, and both %FEV₁/FVC and FVC < LLN for a mixed restrictive and obstructive defect (1).

The measured FEV₁ value derived from the spirometry was expressed as a percentage of predicted FEV₁ value (%FEV₁), which allowed classification of the severity of abnormal respiratory patterns (obstructive, spirometric restrictive, and mixed defects), according to the ATS/ERS

2005 guidelines (18). The %FEV₁ ≥70% of predicted was graded as mild, 60–69% moderate, 50–59% moderately severe, 35–49% severe, and <35% very severe.

Statistical analysis

Differences in the predicted and LLN values of FEV₁, FVC and %FEV₁/FVC across the GLI2012, NHANES III, Knudson, and Siriraj reference equations were analyzed using analysis of variance (ANOVA) with the Bonferroni correction. Differences in the spirometry diagnosis and in the classifying of the severity of impairment between the GLI2012, NHANES III, Knudson, and Siriraj reference equations were analyzed using nominal analysis of variance (NANOVA). Post-hoc tests were analyzed using a Chi-square test and the P values adjusted with the Bonferroni correction with a level of significance =0.008 (P value of 0.05/number of total variables tested =6). Agreement on interpretation of spirometry results between the two

reference equations was analyzed using kappa (κ). The Kappa statistic was interpreted as follows: 0.81–1.00 was regarded as very good, 0.61–0.80 as good, 0.41–0.60 as moderate, 0.21–0.40 as fair, and ≤0.20 poor (19). All statistical procedures were performed using SPSS version 22 (IBM Corp., Armonk, NY, USA). The significance level was set at P<0.05.

Results

A total of 2,863 spirometry results were included in this study. Data from 371 subjects failed to meet the inclusion criteria (*Figure 1*). The total number of spirometry results included in the final analysis was 2,492. The characteristics of the participants are shown in *Table 1*. Spirometry outcomes and predicted values from the GLI2012, NHANES III, Knudson, and Siriraj reference equations are shown in *Tables S1,S2*. The predicted values for FEV₁, FVC, and %FEV₁/FVC were significantly different across four reference equations in both genders (P<0.001). An exception to this was the FVC between the GLI2012 and NHANES III reference equations in men. The highest difference of predicted values for FEV₁ was found when results were compared between the Knudson and GLI2012 reference equations (240 mL) in men and between the Knudson and Siriraj reference equations (110 mL) in women. The highest mean difference of predicted values for FVC was found when comparison between the NHANES III, GLI2012 (250 mL) in men and between the Siriraj and NHANES III reference equations (180 mL) in women.

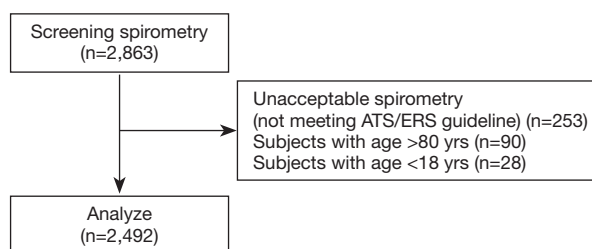


Figure 1 Flow-chart describing the study population.

Table 1 Demographic data (n=2,492)

Variables	Men	Women	Total
Number of subjects	1,336 (53.6)	1,156 (46.4)	2,492
Age (years)	56.16±15.28	53.52±15.05	54.93±15.22
Age categories			
18–25 years	74 (5.5)	63 (5.4)	137 (5.5)
26–39 years	119 (8.9)	152 (13.1)	271 (10.9)
40–59 years	538 (40.3)	523 (45.2)	1,061 (42.6)
60–80 years	605 (45.3)	418 (36.2)	1,023 (41.1)
Height (cm)	162.88±7.01	152.33±6.37	157.99±8.53
Body weight (kg)	59.78±13.59	52.57±11.69	56.44±13.24
BMI (kg/m ²)	22.43±4.43	22.62±4.70	22.52±4.56

Data are presented as mean ± SD or number (percentage). BMI, body mass index.

Table 2 Distribution of the diagnoses using the four reference equations

Spirometry interpretation	Reference equations*			
	Siriraj ^{a,b,c}	NHANES III ^{b,c,d}	Knudson ^{a,c,d}	GLI2012 ^{a,b,d}
Men (n=1,336)				
Normal	452 (33.8)	605 (45.3)	608 (45.5)	570 (42.7)
Restriction	125 (9.4)	265 (19.8)	110 (8.2)	177 (13.2)
Obstruction	560 (41.9)	275 (20.6)	509 (38.1)	397 (29.7)
Mixed pattern	199 (14.9)	191 (14.3)	109 (8.2)	192 (14.4)
Women (n=1,156)				
Normal	421 (36.4)	517 (44.7)	583 (50.4)	521 (45.1)
Restriction	125 (10.8)	279 (24.1)	144 (12.5)	169 (14.6)
Obstruction	449 (38.8)	168 (14.5)	299 (25.9)	300 (26.0)
Mixed pattern	161 (13.9)	192 (16.6)	130 (11.2)	166 (14.4)
Total (n=2,492)				
Normal	873 (35.0)	1,122 (45.0)	1,191 (47.8)	1,091 (43.8)
Restriction	250 (10.0)	544 (21.8)	254 (10.2)	346 (13.9)
Obstruction	1,009 (40.5)	443 (17.8)	808 (32.4)	697 (28.0)
Mixed pattern	360 (14.4)	383 (15.4)	239 (9.6)	358 (14.4)

Data are shown as number (percentage). *P<0.001, compared across the four reference equations using NANOVA test; P<0.008 by Chi-square test, ^acompared with NHANES III, ^bcompared with Knudson, ^ccompared with GLI2012, ^dcompared with Siriraj.

LLN values from the four reference equations are shown in *Tables S1,S2*. The LLN values for FEV₁, FVC, and %FEV₁/FVC were significantly different across four reference equations in both genders (P<0.001). The highest difference of LLN values for FEV₁ was 300 and 80 mL when comparison was made between the Knudson and GLI2012 reference equations in men and between the Knudson and NHANES III reference equations in women, respectively. The highest mean difference of LLN values for FVC was found when comparison was made between the Knudson and NHANES III reference equations in both men and women (470 and 260 mL, respectively). The highest and lowest LLN values of %FEV₁/FVC was found when using the Siriraj and NHANES III reference equations in both genders.

Significant differences in diagnosis spirometry were found across four reference equations in both gender and total subjects (P<0.001) (*Table 2*). The highest number of normal spirometric results was found when using the Knudson, while the lowest rate was found with the Siriraj reference equations in both men and women. The highest rate of airflow obstruction (AO) was found using the Siriraj,

while the lowest rate was found using the NHANES III reference equations in both men and women. Changing from the NHANES III to Siriraj reference equations caused a maximum increase in rates of AO (103.6% and 167.3% in men and women, respectively), while changing from the Knudson to Siriraj and GLI2012 reference equations resulted in a minimum increase in rates of AO (10.1% and 0.3% in men and women, respectively). The highest spirometric restriction rate was found when using the NHANES III in both men and women. Changing from the Knudson and Siriraj to the NHANES III reference equations caused a maximum increase in rates of spirometric restriction (140.9% and 123.2% in men and women, respectively). The lowest rate of mixed defects was found when using the Knudson reference equations in both men and women.

There were significant differences in the classification of severity of respiratory impairment across the four reference equations in both genders (P<0.001) (*Table 3*). The highest number of subjects with mild impairment for abnormal spirometric results was found when using the Siriraj reference equations in both genders.

Table 3 Distribution of the severity classification of abnormal spirometry results using the four reference equations

Severity classification	Reference equations*			
	Siriraj ^{a,b,c}	NHANES III ^{b,c,d}	Knudson ^{a,c,d}	GLI2012 ^{a,b,d}
Men, n (%)	n=884	n=731	n=728	n=766
Mild	423 (47.9)	265 (36.3)	319 (43.8)	221 (28.9)
Moderate	142 (16.1)	147 (20.1)	110 (15.1)	160 (20.9)
Moderately severe	111 (12.6)	119 (16.3)	125 (17.2)	124 (16.2)
Severe	139 (15.7)	128 (17.5)	110 (15.1)	168 (21.9)
Very severe	69 (7.8)	72 (9.8)	64 (8.8)	93 (12.1)
Women, n (%)	n=735	n=639	n=573	n=635
Mild	327 (44.5)	202 (31.6)	137 (23.9)	197 (31.0)
Moderate	131 (17.8)	144 (22.5)	124 (21.6)	140 (22.0)
Moderately severe	109 (14.8)	110 (17.2)	104 (18.2)	108 (17.0)
Severe	115 (15.6)	127 (19.9)	143 (25.0)	129 (20.3)
Very severe	53 (7.2)	56 (8.8)	65 (11.3)	61 (9.6)
Total, n (%)	n=1,619	n=1,370	n=1,301	n=1,401
Mild	750 (46.3)	467 (34.1)	456 (35.0)	418 (29.8)
Moderate	273 (16.9)	291 (21.2)	234 (18.0)	300 (21.4)
Moderately severe	220 (13.6)	229 (16.7)	229 (17.6)	232 (16.6)
Severe	254 (15.7)	255 (18.6)	253 (19.4)	297 (21.2)
Very severe	122 (7.5)	128 (9.3)	129 (9.9)	154 (11.0)

* P<0.001, compared across the four reference equations using NANOVA test; P<0.008 by Chi-square test, ^acompared with NHANES III, ^bcompared with Knudson, ^ccompared with GLI2012, ^dcompared with Siriraj.

The levels of agreement on diagnosis of spirometry results among the four reference equations were found to be moderate to good in both men and women (Kappa values ranged from 0.59 to 0.77 for men and 0.53 to 0.80 for women, respectively) (*Table 4*). The highest kappa value for diagnosis was found between the NHANES III and GLI2012 reference equations in men and the Knudson and GLI2012 reference equations in women. The percentage of agreement of diagnosis was highest when using the NHANES III and GLI2012 in men and the Knudson and GLI2012 reference equations in women. The overall levels of agreement among all four reference equations were moderate to good (Kappa values ranging from 0.56 to 0.77).

The levels of agreement on the classifying severity of respiratory impairment among all four reference equations were found to be moderate to good and good to very good in men and women, respectively (Kappa values ranged from 0.49 to 0.84 for men and 0.71 to 0.90 for

women, respectively) (*Table 5*). The highest kappa value was found between the Siriraj and NHANES III reference equations in men and between the GLI2012 and NHANES III reference equations in women. The percentages of agreement of severity classification were highest using the Siriraj and NHANES III and the GLI2012 and NHANES III reference equations in men and women, respectively. The overall levels of agreement among all four reference equations were moderate to very good (Kappa values ranged from 0.67 to 0.82).

Discussion

This study demonstrated differences in the spirometry interpretation among the four reference equations including all from GLI2012, NHANES III, Knudson, and Siriraj in both genders and total subjects. However, moderate to very good levels of agreement on interpreting spirometry

Table 4 Levels of agreement for diagnosis of spirometry results between the four reference equations

Level of agreement	Siriraj	NHANES III	Knudson
Men (n=1,336)			
Kappa value (95% CI)			
NHANES III	0.59 (0.56–0.62)*	–	–
Knudson	0.70 (0.67–0.73)*	0.59 (0.56–0.62)*	–
GLI2012	0.77 (0.74–0.80)*	0.77 (0.74–0.80)*	0.73 (0.70–0.76)*
% agreement			
Siriraj	–	–	–
NHANES III	70.06	–	–
Knudson	80.01	71.56	–
GLI2012	83.98	83.91	81.89
Women (n=1,156)			
Kappa value (95% CI)			
NHANES III	0.53 (0.50–0.56)*	–	–
Knudson	0.69 (0.66–0.72)*	0.67 (0.64–0.70)*	–
GLI2012	0.76 (0.73–0.79)*	0.72 (0.69–0.75)*	0.80 (0.77–0.83)*
% agreement			
NHANES III	65.92	–	–
Knudson	79.32	77.59	–
GLI2012	83.39	80.54	86.94
Total (n=2,492)			
Kappa value (95% CI)			
NHANES III	0.56 (0.54–0.58)*	–	–
Knudson	0.70 (0.68–0.72)*	0.63 (0.60–0.65)*	–
GLI2012	0.77 (0.75–0.79)*	0.75 (0.73–0.77)*	0.77 (0.74–0.79)*
% agreement			
NHANES III	68.13	–	–
Knudson	79.69	74.36	–
GLI2012	83.70	82.34	84.22

*P<0.001 by Kappa statistic.

results were found among these four reference equations. The results of this study suggest that changing spirometry reference equations would result in different interpretations of spirometry results in Thai people.

The results of this study demonstrated that changing between the four reference equations (the GLI2012, NHANES III, Knudson, and Siriraj reference equations)

resulted in a change in rate of abnormality lung function in both genders. The overall highest rate of abnormality as regards spirometry results was found using the Siriraj reference equations. The results are supported by previous findings showing that changing from the commonly used to another spirometry reference equation resulted in a change in rate of abnormality of lung function (14,15,20–22).

Table 5 Levels of agreement for severity classification of abnormal spirometry results between the four reference equations

Level of agreement	Siriraj	NHANES III	Knudson
Men			
Kappa value (95% CI)			
NHANES III	0.84 (0.81–0.87)*	–	–
Knudson	0.68 (0.64–0.72)*	0.77 (0.73–0.81)*	–
GLI2012	0.62 (0.58–0.66)*	0.60 (0.56–0.64)*	0.49 (0.45–0.53)*
% agreement			
NHANES III	87.56	–	–
Knudson	76.54	82.25	–
GLI2012	70.69	68.64	60.24
Women			
Kappa value (95% CI)			
NHANES III	0.79 (0.75–0.83)*	–	–
Knudson	0.71 (0.67–0.75)*	0.83 (0.80–0.86)*	–
GLI2012	0.82 (0.79–0.85)*	0.90 (0.87–0.93)*	0.89 (0.86–0.92)*
% agreement			
NHANES III	83.79	–	–
Knudson	77.25	86.44	–
GLI2012	86.29	92.78	91.15
Total			
Kappa value (95% CI)			
NHANES III	0.82 (0.79–0.84)*	–	–
Knudson	0.70 (0.67–0.73)*	0.80 (0.77–0.83)*	–
GLI2012	0.71 (0.68–0.74)*	0.74 (0.71–0.77)*	0.67 (0.64–0.70)*
% agreement			
NHANES III	85.82	–	–
Knudson	76.85	84.19	–
GLI2012	77.73	79.66	74.13

*P<0.001 by Kappa statistic; Number of male subjects for agreement between Siriraj and NHANES III =675, Siriraj and Knudson =716, Siriraj and GLI2012 =754, NHANES III and Knudson =614, NHANES III and GLI2012 =676, Knudson and GLI2012 =679.

For example, Collen *et al.* (20) found that changing from the NHANES III to Knudson reference equations caused a 19.5% reduction in the rate of AO. Aggarwal *et al.* (14) also showed that changing from the north Indian to Caucasian reference equations including the NHANES III, Knudson, European Community for Coal and Steel (ECCS), and Crapo caused a change in rate of abnormality in spirometry results in both AO and spirometric restriction. Likewise,

other studies showed that changing from the NHANES III to GLI2012 reference equations resulted in a 25.2% reduction in the rate of spirometric restriction and a 22.1% reduction in the rate of mixed defects (21,22). There is evidence to suggest that spirometry results should be interpreted with caution when the %FEV₁/FVC is close to the LLN values (23). Notably, the differences in rate of AO between the four reference equations in

this study likely occurred when the spirometric results, especially in the case of the %FEV₁/FVC, was close to the LLNs values. Those individuals just falling into the mild abnormality classification in one reference equation would be categorized as having a normal spirometry result using another reference equation.

The classification of the severity of abnormal spirometry results depends on the predicted value of FEV₁ from each reference equation (1) and our study also showed that the predicted values of FEV₁ were significantly different across the four reference equations in both genders. The rates of the severity classification were changed when changing across the four reference equations in both genders, in particular in the case of a mild degree of impairment. Our results are comparable to previous findings demonstrating that the rates of severity classification of AO were changed when changing from the Korean (Lee and Choi) to Caucasian (Morris and Knudson) reference equations (15). Other studies also showed that the rates of severity classification were changed when changing from the NHANES III to the GLI2012 reference equations in a Caucasian population (21,24).

Although the predicted and LLN values for the FVC and FEV₁ estimated using the NHANES III and Knudson reference equations were already adjusted for use in a Thai population, differences in predicted and LLN values across the four reference equations in both genders were observed. It is known that interpretation of spirometry is dependent on predicted values. The differences in predicted and LLN values among the four reference equations may be associated with various factors especially with respect to ethnicity, sample size, and statistical analysis used. As regards the ethnicity factor, the Siriraj reference equation is derived only from a Thai population. The GLI2012 reference equations used for analysis in this study were derived from the Southeast Asian subgroup including Taiwan, South of China, Hong Kong, and Thailand (13), the NHANES III reference equations derived from Caucasian, Mexican-American, and African-American subjects, and the Knudson reference equations derived from the white non-Mexican-American population in America. A multi ethnic group may lead to differences in predicted and LLN values of FEV₁, FVC, and %FEV₁/FVC. There is evidence suggesting that factors associated with ethnicity can contribute to differences in the spirometry predicted values (25). Additionally, the variety of sample size used in each study may influence the predicted and LLN values. In this study it was discovered that the GLI2012 reference

equations were estimated from the largest sample size (n=8,255 for Southeast Asian population), whilst the NHANES III, Siriraj, and Knudson reference equations were estimated from random samples of 7,429, 3,954, and 697, respectively. Variance of sample size may influence predicted and LLN values for FVC, FEV₁, and %FEV₁/FVC. The final factor that may influence the differences in predicted value is the statistical test used. The GLI2012 reference equations were calculated using the lambda-mu-sigma (LMS) method whereas the NHANES III, Knudson, and Siriraj reference equations were estimated from regression analyses. It has been suggested that the LMS statistical method allows modeling of variability and skewedness of data and uses splines to allow for the interactive effects of age, height, and sex (16).

The results of this study showed satisfactory levels of agreement in diagnosis and classification of the severity of respiratory impairment between the GLI2012, NHANES III, Knudson, and Siriraj reference equations. These results are consistent with a previous study adding weight to the finding that there is a good level of agreement in the interpretation of spirometry between the GLI2012, NHANES III, ECCS, and Stanojevic reference equations (Kappa value ranged from 0.72–0.94) (21). A study in an Indian population also showed a moderate to good level of agreement in the interpretation of spirometry results among north Indian and Caucasian prediction equations including NHANES III, ECCS, Crapo, and Knudson (14). Kappa values from the agreement analysis are commonly used to indicate discordance between raters (methods). However, bias is not reflected in kappa values. Thus the results may not be directly interchangeable (26,27).

The results of this study provide valuable clinical information how interpretation of clinical spirometry results change when using different reference equations in a Thai population. This information is useful for clinicians as it raises the awareness of the importance of the interpretation of spirometry testing data in patients with or without respiratory diseases. However, this study has some limitations. Firstly, there was no data from childhood and adolescent subjects aged less than 18 and older adults, aged more than 80 years due to limitations of available data and also limited data from reference values of the NHANES III reference equations. These may limit the generalizability and transferability (external validity) of the study's findings. Secondly, the associations between the interpretation of spirometry results and clinical symptoms were not addressed in this study. Therefore, investigation into the

correlations between clinical symptoms and the differences in spirometry reference equations used should be addressed in future studies.

Conclusions

Changing between the GLI2012, NHANES III, Knudson, and Siriraj reference equations alters the interpretation of spirometry results for the diagnosis and classification of severity of respiratory impairment in this Thai adult population. However, there were moderate to good levels of agreement on the interpretation of spirometry results between these four reference equations. In clinical practice, physicians should be aware that changing between spirometry reference equations can result in a change of interpretation of spirometry results, which may impact the clinical management in some patients with respiratory diseases.

Acknowledgements

The authors would like to give their special thanks to the staff members of the Division of Pulmonary, Critical Care and Allergy, Department of Internal Medicine, Faculty of Medicine, Chiang Mai University, for their contribution to this study. We would also like to thank Mrs. Antika Wongthanee for her excellent advice on statistical analysis.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The study was approved by the Research Ethics Committee of the Faculty of Medicine, Chiang Mai University [Institutional Review Board (IRB), approval number: Med-2559-04324].

References

1. Pellegrino R, Viegi G, Brusasco V, et al. Interpretative strategies for lung function tests. *Eur Respir J* 2005;26:948-68.
2. Sood A, Dawson BK, Henkle JQ, et al. Effect of change of reference standard to NHANES III on interpretation of spirometric 'abnormality'. *Int J Chron Obstruct Pulmon Dis* 2007;2:361-7.
3. Stanojevic S, Wade A, Stocks J. Reference values for lung function: past, present and future. *Eur Respir J* 2010;36:12-9.
4. Miller A. Lung function testing: selection of reference values and interpretative strategies. *Am Rev Respir Dis* 1992;146:1368-9.
5. Rosenfeld M, Pepe MS, Longton G, et al. Effect of choice of reference equation on analysis of pulmonary function in cystic fibrosis patients. *Pediatr Pulmonol* 2001;31:227-37.
6. Subbarao P, Lebecque P, Corey M, et al. Comparison of spirometric reference values. *Pediatr Pulmonol* 2004;37:515-22.
7. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999;159:179-87.
8. Knudson RJ, Lebowitz MD, Holberg CJ, et al. Changes in the normal maximal expiratory flow-volume curve with growth and aging. *Am Rev Respir Dis* 1983;127:725-34.
9. Ghio AJ, Crapo RO, Elliott CG. Reference equations used to predict pulmonary function. Survey at institutions with respiratory disease training programs in the United States and Canada. *Chest* 1990;97:400-3.
10. Dejsomritrurai W, Nana A, Maranetra KN, et al. Reference spirometric values for healthy lifetime nonsmokers in Thailand. *J Med Assoc Thai* 2000;83:457-66.
11. Pothirat C, Chaiwong W, Phetsuk N, et al. A comparative study of COPD burden between urban vs rural communities in northern Thailand. *Int J Chron Obstruct Pulmon Dis* 2015;10:1035-42.
12. Pothirat C, Chaiwong W, Phetsuk N, et al. Long-term efficacy of intensive cycle ergometer exercise training program for advanced COPD patients. *Int J Chron Obstruct Pulmon Dis* 2015;10:133-44.
13. Quanjer PH, Stanojevic S, Cole TJ, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012;40:1324-43.
14. Aggarwal AN, Gupta D, Behera D, et al. Applicability of commonly used Caucasian prediction equations for spirometry interpretation in India. *Indian J Med Res* 2005;122:153-64.
15. Kim N, Park MH, Kim SY, et al. Discordance in spirometric interpretations based on korean and non-korean reference equations. *Ann Occup Environ Med* 2013;25:42.
16. Quanjer PH, Hall GL, Stanojevic S, et al. Age- and height-based prediction bias in spirometry reference equations. *Eur Respir J* 2012;40:190-7.
17. Pereira CA, Duarte AA, Gimenez A, et al. Comparison

- between reference values for FVC, FEV1, and FEV1/FVC ratio in White adults in Brazil and those suggested by the Global Lung Function Initiative 2012. *J Bras Pneumol* 2014;40:397-402.
18. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J* 2005;26:319-38.
 19. Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991.
 20. Collen J, Greenburg D, Holley A, et al. Discordance in spirometric interpretations using three commonly used reference equations vs national health and nutrition examination study III. *Chest* 2008;134:1009-16.
 21. Brazzale DJ, Hall GL, Pretto JJ. Effects of adopting the new global lung function initiative 2012 reference equations on the interpretation of spirometry. *Respiration* 2013;86:183-9.
 22. Quanjer PH, Brazzale DJ, Boros PW, et al. Implications of adopting the Global Lungs Initiative 2012 all-age reference equations for spirometry. *Eur Respir J* 2013;42:1046-54.
 23. Culver BH. How should the lower limit of the normal range be defined? *Respir Care* 2012;57:136-45.
 24. Embling LA, Zagami D, Sriram KB, et al. Effect of changing from the National Health and Nutritional Examination Survey III spirometry reference range to that of the Global Lung Initiative 2012 at Gold Coast Hospital and Health Service. *J Thorac Dis* 2016;8:3739-43.
 25. Lung function testing: selection of reference values and interpretative strategies. American Thoracic Society. *Am Rev Respir Dis* 1991;144:1202-18.
 26. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22:276-82.
 27. Hoehler FK. Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. *J Clin Epidemiol* 2000;53:499-503.

Cite this article as: Chaiwong W, Uthairakun S, Liwsrisakun C, Pothirat C. Effect of changing reference equations for spirometry interpretation in Thai people. *J Thorac Dis* 2019;11(1):113-122. doi: 10.21037/jtd.2018.11.06

Table S1 Spirometry outcomes and predicted values calculated from the four reference equations

Parameters	Measured spirometric data	Mean predicted value			
		Siriraj	NHANES III	Knudson	GLI2012
Men (n=1,336)					
FEV ₁ (L)*	2.09±0.85	2.60±0.54 ^{a,b,c}	2.63±0.59 ^{b,c,d}	2.56±0.66 ^{a,c,d}	2.80±0.55 ^{a,b,d}
FVC (L)*	2.94±0.91	3.25±0.60 ^{a,b,c}	3.41±0.63 ^{b,d}	3.16±0.73 ^{a,c,d}	3.41±0.59 ^{b,d}
%FEV ₁ /FVC*	69.61±15.10	83.44±2.59 ^{a,b,c}	76.47±3.15 ^{b,c,d}	81.19±1.91 ^{a,c,d}	81.63±2.64 ^{a,b,d}
Women (n=1,156)					
FEV ₁ (L)*	1.57±0.64	1.98±0.39 ^{a,b,c}	2.04±0.43 ^{b,c,d}	2.09±0.39 ^{a,c,d}	2.06±0.40 ^{a,b,d}
FVC (L)*	2.11±0.68	2.34±0.41 ^{a,b,c}	2.56±0.44 ^{b,c,d}	2.52±0.42 ^{a,c,d}	2.46±0.42 ^{a,b,d}
%FEV ₁ /FVC*	73.35±13.36	85.29±1.64 ^{a,b,c}	79.46±3.16 ^{b,c,d}	83.58±2.60 ^{a,c,d}	83.49±2.65 ^{a,b,d}
Total (n=2,492)					
FEV ₁ (L)*	1.85±0.82	2.31±0.57 ^{a,b,c}	2.35±0.60 ^{b,c,d}	2.34±0.60 ^{a,c,d}	2.45±0.6 ^{a,b,d}
FVC (L)*	2.55±0.91	2.83±0.69 ^{a,b,c}	3.01±0.69 ^{b,c,d}	2.86±0.68 ^{a,c,d}	2.97±0.70 ^{a,b,d}
%FEV ₁ /FVC*	71.34±14.43	84.30±2.39 ^{a,b,c}	77.85±3.49 ^{b,c,d}	82.30±2.55 ^{a,c,d}	82.50±2.80 ^{a,b,d}

Data are presented as mean ± SD. *P<0.001, compared across the NHANES III, Knudson, GLI2012, and Siriraj reference equations by ANOVA; P<0.05, ^acompared with NHANES III, ^bcompared with Knudson, ^ccompared with GLI2012, ^dcompared with Siriraj. FEV₁, forced expiratory volume in the first second; FVC, forced vital capacity.

Table S2 Lower limit of normal values calculated from the four reference equations

Parameters	Mean lower limit of normal value			
	Siriraj	NHANES III	Knudson	GLI2012
Men (n=1,336)				
FEV ₁ (L)*	1.99±0.55 ^{a,b,c}	2.04±0.56 ^{b,c,d}	1.81±0.62 ^{a,c,d}	2.11±0.50 ^{a,b,d}
FVC (L)*	2.54±0.60 ^{a,b,c}	2.73±0.58 ^{b,c,d}	2.26±0.69 ^{a,c,d}	2.61±0.53 ^{a,b,d}
%FEV ₁ /FVC*	74.65±2.54 ^{a,b,c}	66.82±3.21 ^{b,c,d}	71.57±2.52 ^{a,c,d}	70.83±3.52 ^{a,b,d}
Women (n=1,156)				
FEV ₁ (L)*	1.53±0.39 ^{a,b,c}	1.59±0.39 ^{b,c,d}	1.51±0.37 ^{a,c,d}	1.55±0.36 ^{a,b,d}
FVC (L)*	1.81±0.41 ^{a,b,c}	2.03±0.40 ^{b,c,d}	1.77±0.39 ^{a,c,d}	1.87±0.38 ^{a,b,d}
%FEV ₁ /FVC*	77.09±1.63 ^{a,b,c}	69.68±3.17 ^{b,c,d}	71.52±2.87 ^{a,c,d}	72.99±3.30 ^{a,b,d}
Total (n=2,492)				
FEV ₁ (L)*	1.78±0.53 ^{a,b,c}	1.83±0.54 ^{b,c,d}	1.67±0.54 ^{a,c,d}	1.85±0.53 ^{a,b,d}
FVC (L)*	2.20±0.64 ^{a,b,c}	2.40±0.62 ^{b,c,d}	2.03±0.62 ^{a,c,d}	2.26±0.59 ^{a,b,d}
%FEV ₁ /FVC*	75.79±2.49 ^{a,b,c}	68.15±3.49 ^{b,c,d}	71.50±2.69 ^{a,c,d}	71.83±3.59 ^{a,b,d}

Data are presented as mean ± SD. *P<0.001, compared across the NHANES III, Knudson, GLI2012, and Siriraj reference equations by ANOVA; P<0.05, ^acompared with NHANES III, ^bcompared with Knudson, ^ccompared with GLI2012, ^dcompared with Siriraj. FEV₁, forced expiratory volume in the first second; FVC, forced vital capacity.