Original Article

A New Derivation of the Reference Equations of Spirometric Values in the Thai Population Using the GAMLSS Method

Wanchai Dejsomritrutai MD, MSc[1], Benjamas Chuaychoo MD, PhD[1], Nitipatana Chierakul MD[1]

[1] Division of Respiratory Disease and Tuberculosis, Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand

Objective: The Global Lung Initiative (GLI)-2012 equation has now been used as the reference for the interpretation of the spirometry worldwide. The equations have been derived using modern statistical methods called Generalized Additive Models for Location, Scale and Shape (GAMLSS) and Spline. A part of the spirometric data from which the equations were derived is the data collected from the Thai population. The present study aimed to create the reference equations using only the data from the Thai population retrieved from the previous cross-sectional study and compared them to the GLI-2012's.

Materials and Methods: The previous data of each spirometric parameters including FVC, FEV_1, FEV_1/FVC, and $FEF_{25-75\%}$ in 1,616 male and 2,280 female subjects were retrieved for the statistical modeling by GAMLSS and Spline packages in R Statistics. Lambda-Mu-Sigma (LMS) models were used to establish the best fitted models defined by the smallest Schwarz's Bayesian Criterion (SBC). The authors also used a set of spirometric data of healthy elderly subjects to validate the equations and compare with the GLI-2012.

Results: The GAMLSS-Thai equations were then derived to establish the predicted (Mu) and the lower limit of normal (LLN) values for each parameter. For 140 male and 651 female subjects aged 60 to 80 years, the predicted performance of GAMLSS-Thai is better than GLI-2012 equations as shown by various statistical tools including z-score, mean absolute percentage error (MAPE), and root mean square deviation (RMSD) in all spirometric parameters.

Conclusion: The GAMLSS-Thai equation can be alternatively used as reference spirometric values in subjects aged 10 to 80 years in Thai population. It is more appropriate for subjects aged 60 to 80 years than the GLI-2012 equation.

Keywords: Spirometry; Reference; Thailand; Population; GAMLSS

Received 13 October 2021 | Revised 14 December 2021 | Accepted 14 December 2021

J Med Assoc Thai 2022;105(1):15-21

Website: http://www.jmatonline.com

The interpretation of spirometry requires a comparison between measured values with the standard or predicted ones. The values below or above the lower limit of normal (LLN) or upper limit of normal, respectively, are considered abnormal. Besides, the predicted, as well as the LLN, are dependent on substantial variables including gender, age, height, and ethnicity. The equations of the predicted spirometric values have thus been derived from the survey studies among general populations usually in each ethnic background.

A nationwide survey for spirometric values among the never-smoker population in Thailand was conducted and the reference equations for spirometry were accordingly derived(1). The equations have been widely used as the reference values in Thai patients. Nevertheless, the derivation of equations was established by multiple regression analysis considered as the most proper statistical method at that time. However, there have been major problems recognized of the method including the problem of the transitional zone between childhood and adult(2) and the difficulty for establishing the equations for the LLN.

In 2012, the Global Lung Initiative (GLI), a collaboration group of international academics has retrieved spirometric data worldwide, including the data from Thailand as mentioned, and has published a set of reference equations of spirometry derived from modern statistical methods called Generalized

Correspondence to:

Dejsomritrutai W.

Division of Respiratory Disease and Tuberculosis, Department of Medicine, Faculty of Medicine Siriraj Hospital, 2 Wanglang Road, Bangkoknoi, Bangkok 10700, Thailand.

Phone: +66-2-4197757, **Fax**: +66-2-4197760

Email: wanchai.dej@mahidol.ac.th

How to cite this article:

Dejsomritrutai W, Chuaychoo B, Chierakul N. A New Derivation of the Reference Equations of Spirometric Values in the Thai Population Using the GAMLSS Method. J Med Assoc Thai 2022;105:15-21.

DOI: 10.35755/jmedassocthai.2022.01.13227

Additive Models for Location, Scale and Shape (GAMLSS) as well as spline smoothing method[3]. The equations can be used in various ethnic populations including Caucasian, African American, North-east Asian, and South-east Asian, which include the populations in Thailand, Taiwan, and China below Huaihe river and Qinling mountains[3]. The GLI-2012 equations also include the LLN of each spirometry parameter.

Forty to fifty percent of the data for the GLI-2012 references of the southeast Asian was recruited from Thailand[1,3]. Although differences in predicted values between the previous Thai equations[1] and those of GLI-2012[3] are observed especially in extreme age groups, it is doubtful whether the equations derived by the modern statistical method from the data of only Thai population is more accurate.

The new reference equations are then derived using the GAMLSS method and spline smoothing. The spirometric reference values as well as the LLN from the derived equations are compared with those from GLI-2012. The values in subjects with age 60 to 80 years have also been validated.

Materials and Methods

Data collection, described in detail in the previous study[1], was done in a nation-wide survey over five geographical regions in the Thai population with age from 10 years and over. The number of subjects was stratified by the geographical regions. Subjects recruited in the present study were lifetime non-smokers who had no cardio-respiratory symptoms, unremarkable physical examination, and normal chest radiography. For the present analysis, data retrieved from the previous study were gender, age, height, and spirometric parameters including forced vital capacity (FVC), forced expiratory volume in one second (FEV_1), FEV_1/FVC, and forced expiratory flow 25% to 75% ($FEF_{25-75\%}$). The study has been approved by Siriraj Institutional Review board (COA no. Si 459/2020).

Statistical methods

The R statistical software version 4.0.2 was used for the analysis. The modeling was performed by generalized additive models for location, scale, and shape (GAMLSS) packages with Spline package. The analysis was done separately for male and female data. The Box-Cox-Cole-Green (BCCG) distribution or Lambda-Mu-Sigma (LMS) method was applied in the GAMLSS method[2,4]. The LMS method allowed for the modeling of predicted value or mu (M), coefficient of variation or sigma (S), and skewness or lambda (L). Additionally, the smooth continuation of the data over the entire age range was performed by spline package functions in R-statistics. The appropriate models in terms of fitting and the degree of freedom were selected by visualized worm plots as well as the smallest Schwarz's Bayesian Criterion (SBC)[4]. The general model for each parameter was shown as the following equation. Where Y is accordingly L, M, or S; a, b, and c were coefficients; spline was age-specific coefficient from spline function; LLN, the LLN was equivalent to the fifth percentile value.

$$Y = e[a*\ln(age) + b*\ln(height) + c + Y\text{-spline}]$$

$$LLN = e[\ln(1-1.644*L*S)/L) + \ln(M)]$$

Validation of the equations

To compare the equations of GAMLSS-Thai with those of GLI-2012 for the Southeast Asian, the data previously retrieved from a respiratory health survey in the older persons in Bangkok was used[5]. The survey was done in 2000 among subjects aged 60 years and over. Only the data of subjects represented as healthy was selected by excluding those with a history of respiratory symptoms, smoking, and abnormal chest radiography. The spirometric parameters of 140 male and 651 female subjects were used for the validation of the equations.

The comparison between GAMLSS-Thai and GLI-2012 equations was done by various statistical methods. The z-scores were calculated as the formula: z-score = [(observed value – predicted value)/ standard deviation (SD)] and the student's t-test was used to determine the difference between the mean and zero. Bland-Altman plots[6] were also used for the comparison of predicted spirometric values of GAMLSS-Thai against the GLI-2012. Mean absolute percentage error (MAPE), as well as root mean square deviation (RMSD) were calculated for the comparison between both regression models.

Results

GAMLSS-Thai equations

The characteristics of subjects whose data were used for the derivation of the reference spirometric equations by GAMLSS statistics for GAMLSS-Thai are shown in Table 1. As an example, the distribution of FEV_1 of male and female subjects by age and height is shown in Figure 1. GAMLSS and Spline package in R statistics were used to create models best fitted to the data. The Box-Cox-Cole-Green (BCCG) distribution was the most appropriate

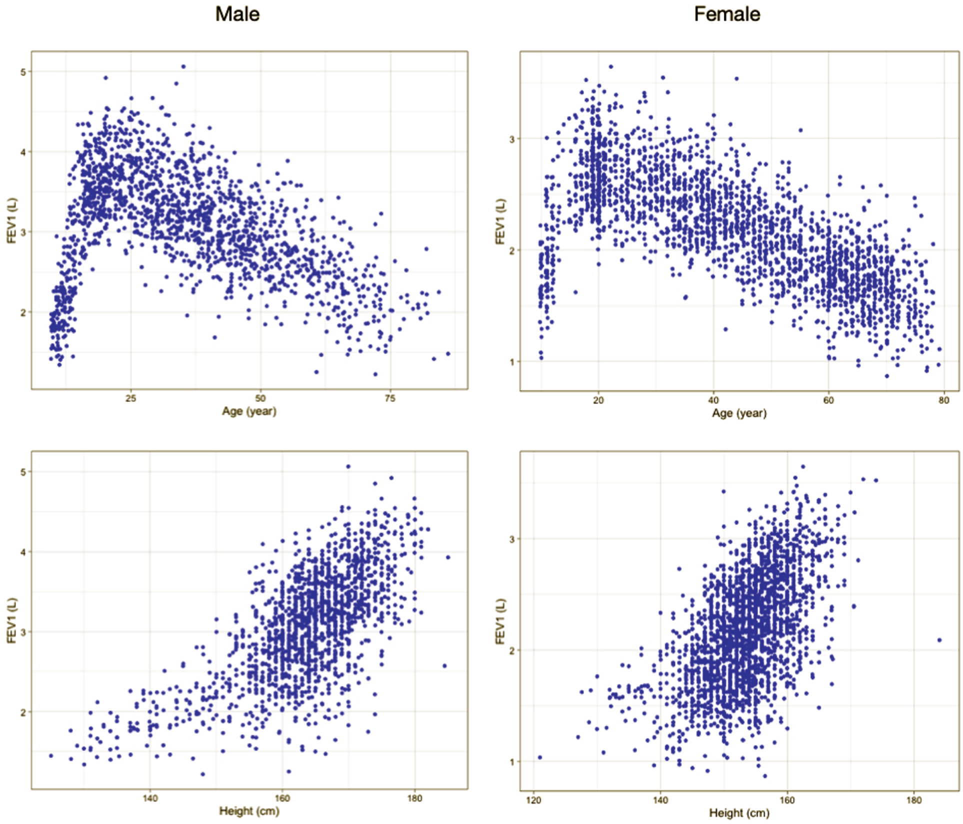

Figure 1. Distributions of FEV1 in the nation-wide survey in Thai population.

Table 1. Demographic data of all subjects and spirometric parameters used for the derivation of the equations

	Female (n=2,280); mean±SD)	Male (n=1,616); mean±SD
Age (year)	42.93±18.76	35.38±17.38
Height (cm)	153.5±6.3	163.6±8.9
FVC (L)	2.509±0.521	3.585±0.748
FEV_1 (L)	2.179±0.484	3.070±0.672
FEV_1/FVC (%)	86.94±5.61	85.65±6.12
$FEF_{25-75\%}$ (L/second)	2.761±0.894	3.609±1.154

SD=standard deviation; FVC=forced vital capacity; FEV_1=forced expiratory volume in 1 second; $FEF_{25-75\%}$=forced expiratory flow 25% to 75%

model according to the smallest SBC and visualizing worm plots. The GAMLSS-Thai equations for the predicted values of M, mu, and LLN were created from the models as shown in Table 2. Figure 2 demonstrates the graphic presentation of predicted and LLN values of each spirometric parameter from both equations as applied to males and females with the height of 168 cm and 159 cm, respectively, within the age range of 20 to 80 years old.

GAMLSS-Thai v GLI-2012 in the older subjects

It has been noted from practical experiences that the spirometry interpretation is less reliable in subjects with extreme ages such as in childhood and elderly. The authors, therefore, retrieved a subset of the data from the previous respiratory health survey among the older persons in Bangkok. Various statistical methods were then used to compare the performance between GAMLSS-Thai and GLI-2012 equations. Table 3 reveals the statistical values used to determine the offsets of calculated predicted values and the validating data including z-scores, MAPE, and RMSD.

GAMLSS-Thai yielded more proper z-scores near zero than GLI-2012 in most spirometric parameters including FEV_1/FVC. Other statistical methods including MAPE and RMSD, which were used to measure the quality for regression models, also provided consistent results. As example, the comparison of z-scores for FEV_1 and FEV_1/FVC are shown in Figure 3. Bland-Altman plots, which quantified the agreement between the two quantitative measurements, also revealed the differences of the

Table 2. The equations for the predicted value (M) and lower limit of normal (LLN) of spirometric parameters derived by GAMLSS statistics

Parameters	Equations		
	Predicted values (M)	S	L
FEV_1			
Male	exp(−9.1244 + 2.0920*logH − 0.1258*logA + M spline)	exp(−2.9714 + 0.2376*logA + S spline)	1
Female	exp(−7.8683 + 1.8780*logH − 0.2256*logA + M spline)	exp(−2.9826 + 0.2475*logA + S spline)	1
FVC			
Male	exp(−10.10021 + 2.27582*logH − 0.06886*logA + M spline)	exp(−2.7898 + 0.1873*logA + S spline)	1
Female	exp(−8.7094 + 2.03*logH − 0.1652*logA + M spline)	exp(−2.7887 + 0.1985*logA + S spline)	1
FEV_1/FVC			
Male	exp(5.41291 − 0.14975*logH − 0.05688*logA + M spline)	exp(−3.2459 + 0.1183*logA + S spline)	2.3474 + 0.3245*logA + L spline
Female	exp(5.30384 − 0.12550*logH − 0.05582*logA + M spline)	exp(−3.7442 + 0.2261*logA + S spline)	7.961 − 1.136*logA + L spline
$FEF_{25-75\%}$			
Male	exp(−5.1923 + 1.4134*logH − 0.2235*logA + M spline)	exp(−2.365 + 0.277*logA + S spline)	0.31447 + 0.03175*logA + L spline
Female	exp(−3.3425 + 1.1072*logH − 0.3436*logA + M spline)	exp(−2.5131 + 0.3047*logA + S spline)	0.74228 − 0.02932*logA + L spline
Lower limit of normal (fifth percentile)	exp(log(1−1.644*L*S)/L + logM))		

FVC=forced vital capacity; FEV_1=forced expiratory volume in 1 second; $FEF_{25-75\%}$=forced expiratory flow 25% to 75%; A=age (year); H=height (cm); L=skewness; M=mean; S=coefficient of variance; log=log_{10}

Table 3. Comparisons of reference spirometric values between GLI-2012 and GAMLSS-Thai equations in Thai subjects aged 60 to 80 years

Statistic tools	Spirometric parameters	Male (n=140)		Female (n=651)	
		GLI-2012	GAMLSS-Thai	GLI-2012	GAMLSS-Thai
z-score (95% CI)	FEV_1	−0.548* (−0.686 to −0.410)	−0.185* (−0.322 to −0.047)	−0.483* (−0.55 to −0.416)	−0.304* (−0.370 to −0.237)
	FVC	−0.719* (−0.854 to −0.583)	−0.325* (−0.460 to −0.190)	−0.725* (−0.792 to −0.658)	−0.349* (−0.414 to −0.285)
	FEV_1/FVC	0.567* (0.407 to 0.727)	0.281* (0.120 to 0.441)	0.638* (0.563 to 0.713)	−0.084* (−0.160 to −0.009)
	$FEF_{25-75\%}$	0.3* (0.182 to 0.491)	0.193* (0.039 to 0.347)	0.332* (0.260 to 0.404)	−0.061 (−0.134 to 0.011)
Proportion below LLN	FEV_1	11.43	9.29	7.22	7.07
	FVC	16.43	13.57	11.98	7.68
	FEV_1/FVC	0.00	0.71	0.00	4.15
	$FEF_{25-75\%}$	2.14	3.14	1.69	5.83
MAPE	FEV_1	17.43	14.0	16.30	14.51
	FVC	19.79	15.29	18.64	14.51
	FEV_1/FVC	5.29	4.73	5.37	4.45
	$FEF_{25-75\%}$	24.87	24.95	24.89	27.25
RMSD	FEV_1	2.79	0.94	7.82	4.91
	FVC	4.50	2.04	13.78	6.64
	FEV_1/FVC	32.54	16.11	169.57	22.40
	$FEF_{25-75\%}$	3.10	1.77	10.7	1.98

GLI=Global Lung Initiative; GAMLSS=Generalized Additive Models for Location, Scale and Shape; CI=confidence interval; LLN=lower limit of normal; MAPE=mean absolute percentage error; RMSD=root mean square deviation; FVC=forced vital capacity; FEV_1=forced expiratory volume in 1 second; $FEF_{25-75\%}$=forced expiratory flow 25% to 75%

* p<0.05, Student's t-test

predicted values between both equations as shown in Figure 4. Note that some points lie beyond 2 standard deviations of mean differences shown as dotted lines. RMSD, a measure of accuracy to compare different models and the value close to 0, was considered a better measure. It was notable that the differences of RMSD in FEV_1/FVC and $FEF_{25-75\%}$ were prominent in the extremely high-aged group, especially in the female (Table 3). Figure 5 also demonstrates the comparisons of the predicted and LLN values between the GLI-2012 and GAMLSS-Thai models in a group of Thai subjects aged 60 to 80 years.

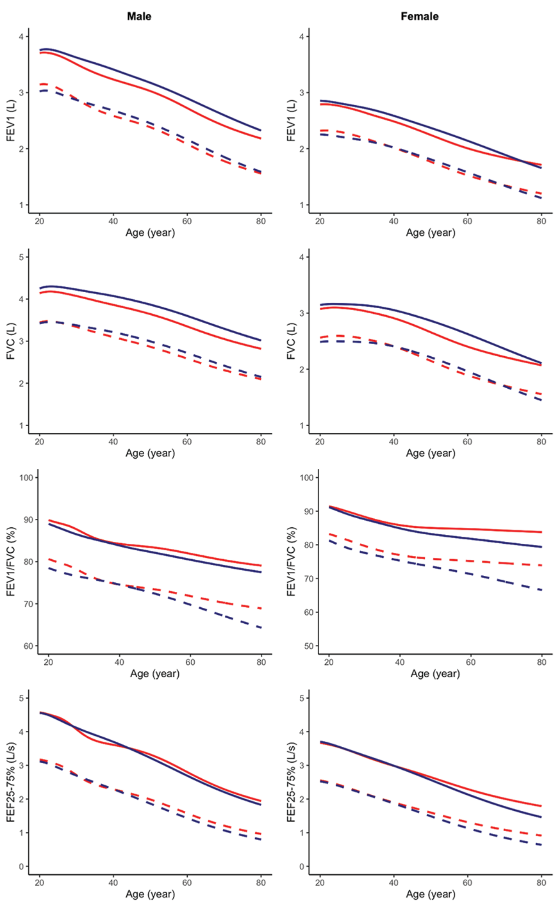

Figure 2. Demonstration of the predicted (continuous line) and lower limit of normal (dash line) values of each spirometric parameter from GLI-2012 (blue) and GAMLSS-Thai (red) equations over age-range of 20 to 80 years in male and female with height of 168 and 159 cm, respectively.

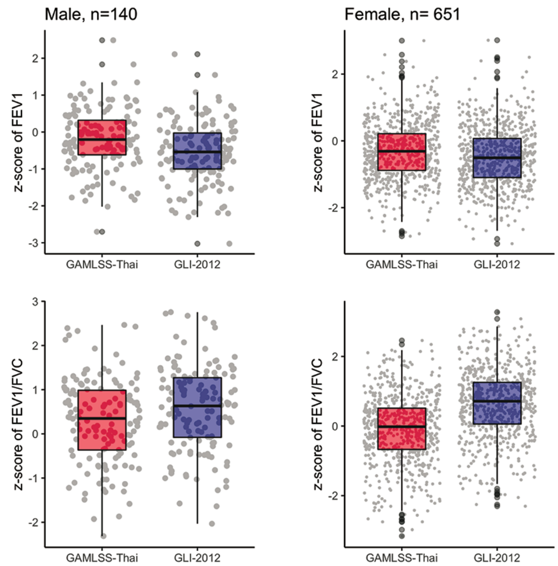

Figure 3. Comparisons of the z-scores of FEV_1 (upper) and FEV_1/FVC (lower) between the GLI-2012 model and GAMLSS-Thai model using a group of Thai subjects aged 60 to 80 years (grey dots) for the validation.

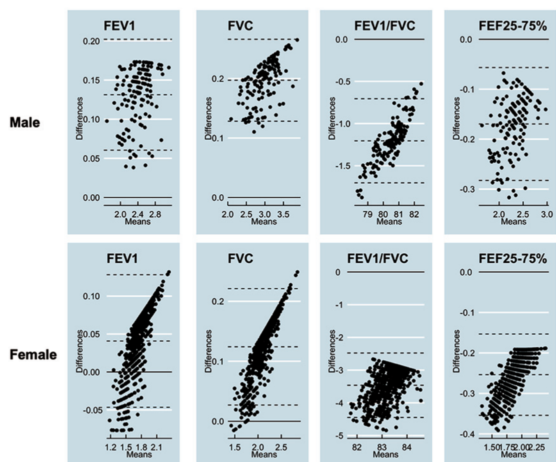

Figure 4. Bland-Altman plots for the comparison of predicted values from GLI-2012 and GAMLSS-Thai in a group of subjects with age 60 to 80 years.

Discussion

The GLI-2012 equation has advantages over the previous spirometric references. The structure of equations can be applied to a population with various ethnicities[3]. The use of modern statistical methods, GAMLSS and Spline, results in the smooth transition of the spirometric values from childhood to adult and the opportunity to determine the LLN values defined as the fifth percentile. The equations, however, were derived from collective data of subjects in several countries across multiple regions. After years of the implementation, it has been questioned whether the equations are suitable to be used in each country. There were reports of the inappropriateness to use the GLI-2012 equations in their population. Jian et al[7] recently developed reference values for spirometry in Chinese aged 4 to 80 years and found that uses of GLI-2012 equations result in significant discrepancies comparing to the newly derived equations. Jo et al[8] also developed the reference spirometric equations by LMS method in Korean adults and reported partial discrepancies between the new reference values from the GLI-2012. There has also been substantial evidence in other regions for the discrepancy[9-11].

In Thailand, the previous Siriraj equations[1] were recommended by the Thoracic Society of Thailand as the spirometric references. However, the GLI-2012 equation is later increasingly used because

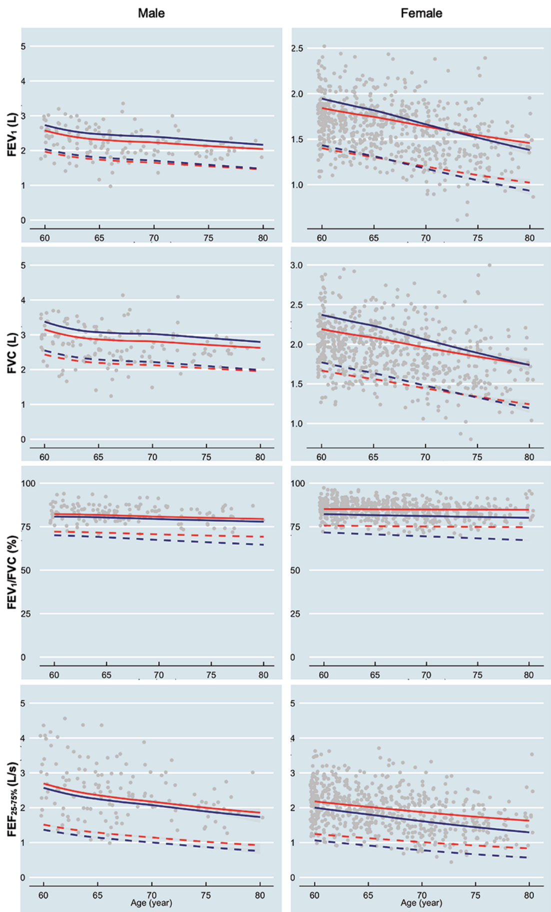

Figure 5. Comparisons of the predicted (continuous line) and lower limit of normal (dash line) values between the GLI-2012 (blue) and GAMLSS-Thai (red) models in a group of Thai subjects aged 60 to 80 years (grey dots).

it provides the LLN values and is supported by international organizations. Although the spirometric data of Thai subjects is already included for the derivation of GLI-2012 equations, it remains in doubt if the authors should derive the equations using the GAMLSS method that is specific to the Thai population, and if it would be more suitable than the GLI-2012.

The GAMLSS-Thai equation is thus derived using the data from the previous survey. The statistical models are developed by GAMLSS and spline methods and the equations for the predicted and LLN values are derived. As compared with the GLI-2012 equation, the predicted values of FEV_1 and FVC from the GAMLSS-Thai equation are lower but the LLN values comparable (Figure 2). However, the major differences are observed in the predicted and LLN values for FEV_1/FVC and $FEF_{25-75\%}$ for which the GAMLSS-Thai yields higher results especially in extremely high-aged subjects (Figure 2).

For the validation, the data from the respiratory health survey in healthy subjects aged 60 years and over are used for the comparison between both equations. Multiple statistical methods including z-scores, MAPE, and RMSD reveal that the GAMLSS-Thai equation results are less error than the GLI-2012 (Table 3, Figure 4). In consideration of FEV_1/FVC, Table 3, also shows that the proportion of subjects with their values less than LLN in GLI-2012 is 0% in both genders compared to 0.71% in male and 4.75% in female in GAMLSS-Thai. Since the LLN value in spirometry is traditionally defined as the fifth percentile value, the GAMLSS-Thai equation provides a more proper value. In other words, the use of the GLI-2012 may be less sensitive to detect the abnormalities as compared to the GAMLSS-Thai.

The reasons for the lower values of FEV_1/FVC and $FEF_{25-75\%}$ in the GLI-2012 equation are due to the inhomogeneous characteristics among subjects. Although the data was collected from countries within the same region, the differences in factors other than biological factors may contribute to the differences in the spirometric values between both equations. Other environmental, socioeconomic, or undetermined factors may also contribute to the discrepancy between both equations. Quanjer et al[12] reported that no secular trends were observed from the analyses of the data collected from several regions, especially in the data retrieved from societies with stable socioeconomic conditions. At present, it is unlikely that secular trends contribute to the differences in pulmonary functions.

Additionally, there are also substantial reports about the variations in the reference spirometric values in the elderly subjects. Healthy subject bias should be in consideration because the retrieved data are more likely collected from more healthy people. The predicted and the LLN values in the older subjects may, thus, be higher than true values. In addition, small sample size in this age group in most studies is another limitation. The spirometry in the elderly should thus be interpreted carefully.

Since the data of the Thai population were retrieved from subjects aged 10 to 80 years, the disadvantage of the GAMLSS-Thai equation is that it cannot be used in children under 10 years. The GLI-2012 equation is thus recommended in children 4 years and older. In contrast, the results of the present study suggest that the GAMLSS-Thai equation is more appropriate in subjects aged 60 to 80 years.

However, further research providing consistent results are needed.

Conclusion

The GAMLSS-Thai equations have been derived using the previous spirometric data from the population in Thailand. The equations can be used alternatively as references for the interpretation of spirometry in the Thai population aged 10 to 80 years. It is more appropriate than GLI-2012 equations in subjects aged 60 to 80 years.

What is already known on this topic?

GLI-2012 equations derived by GAMLSS have been widely used as spirometric references. The data of the Thai population was partly included for generating the equations.

What does this study add?

The GAMLSS-Thai equations are derived only from the data of the population in Thailand and this study demonstrates their utilities as alternative reference equations in Thai people.

Acknowledgement

The authors thank the Ministry of Public Health and the Ministry of Labor and Social Welfare for the support during the previous nationwide survey.

Conflicts of interest

The authors declare no conflict of interest regarding the publication of this paper.

References

1. Dejsomritrutai W, Nana A, Maranetra KN, Chuaychoo B, Maneechotesuwan K, Wongsurakiat P, et al. Reference spirometric values for healthy lifetime nonsmokers in Thailand. J Med Assoc Thai 2000;83:457-66.
2. Stanojevic S, Wade A, Stocks J, Hankinson J, Coates AL, Pan H, et al. Reference ranges for spirometry across all ages: a new approach. Am J Respir Crit Care Med 2008;177:253-60.
3. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, 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.
4. Cole TJ, Stanojevic S, Stocks J, Coates AL, Hankinson JL, Wade AM. Age- and size-related reference ranges: a case study of spirometry through childhood and adulthood. Stat Med 2009;28:880-98.
5. Maranetra KN, Chuaychoo B, Dejsomritrutai W, Chierakul N, Nana A, Lertakyamanee J, et al. The prevalence and incidence of COPD among urban older persons of Bangkok Metropolis. J Med Assoc Thai 2002;85:1147-55.
6. Altman DG, Bland JM. Measurement in Medicine: The Analysis of Method Comparison Studies. J R Stat Soc Ser A Stat Soc 1983;32:307-17.
7. Jian W, Gao Y, Hao C, Wang N, Ai T, Liu C, et al. Reference values for spirometry in Chinese aged 4-80 years. J Thorac Dis 2017;9:4538-49.
8. Jo BS, Myong JP, Rhee CK, Yoon HK, Koo JW, Kim HR. Reference values for spirometry derived using Lambda, Mu, Sigma (LMS) method in Korean adults: in comparison with previous references. J Korean Med Sci 2018;33:e16.
9. Agarwal D, Parker RA, Pinnock H, Roy S, Ghorpade D, Salvi S, et al. Normal spirometry predictive values for the Western Indian adult population. Eur Respir J 2020;56:1902129.
10. Kainu A, Timonen KL, Toikka J, Qaiser B, Pitkäniemi J, Kotaniemi JT, et al. Reference values of spirometry for Finnish adults. Clin Physiol Funct Imaging 2016;36:346-58.
11. Tian XY, Liu CH, Wang DX, Ji XL, Shi H, Zheng CY, et al. Spirometric reference equations for elderly Chinese in Jinan aged 60-84 years. Chin Med J (Engl) 2018;131:1016-22.
12. Quanjer PH, Stocks J, Cole TJ, Hall GL, Stanojevic S. Influence of secular trends and sample size on reference equations for lung function tests. Eur Respir J 2011;37:658-64.