J Korean Med Sci. 2024 Jan 29;39(4):e20
https://doi.org/10.3346/jkms.2024.39.e20
eISSN 1598-6357·pISSN 1011-8934

JKMS

Original Article
Respiratory Diseases & Critical Care Medicine

Spirometric Interpretation and Clinical Relevance According to Different Reference Equations

Hyun Woo Lee [1], **Jung-Kyu Lee** [1], **Yong-Il Hwang** [2], **Hyewon Seo** [3], **June Hong Ahn** [4], **Sung-Ryeol Kim** [5], **Hyun Jung Kim** [6], **Ki-Suck Jung** [7], **Kwang Ha Yoo** [8], **and Deog Kyeom Kim** [1]

[1]Division of Respiratory and Critical Care, Department of Internal Medicine, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul National University College of Medicine, Seoul, Korea
[2]Division of Pulmonary, Allergy and Critical Care Medicine, Department of Internal Medicine, Hallym University Sacred Heart Hospital, Anyang, Korea
[3]Department of Internal Medicine, School of Medicine, Kyungpook National University, Daegu, Korea
[4]Department of Internal Medicine, Yeungnam University Medical Center, College of Medicine, Yeungnam University, Daegu, Korea
[5]Department of Internal Medicine, Yongin Severance Hospital, Yonsei University College of Medicine, Yongin, Korea
[6]Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Keimyung University Dongsan Hospital, Daegu, Korea
[7]Division of Pulmonary, Allergy and Critical Care Medicine, Hallym University Sacred Heart Hospital, Hallym University College of Medicine, Anyang, Korea
[8]Department of Internal Medicine, Konkuk University School of Medicine, Seoul, Korea

Received: Aug 10, 2023
Accepted: Nov 21, 2023
Published online: Dec 20, 2023

Address for Correspondence:
Deog Kyeom Kim, MD
Division of Respiratory and Critical Care, Department of Internal Medicine, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul National University College of Medicine, 20 Boramae-ro-5-gil, Dongjak-gu, Seoul 07061, Korea.
Email: kimdkmd@gmail.com

ORCID iDs
Hyun Woo Lee
https://orcid.org/0000-0003-4379-0260
Jung-Kyu Lee
https://orcid.org/0000-0001-5060-7255
Yong-Il Hwang
https://orcid.org/0000-0002-5777-7983
Hyewon Seo
https://orcid.org/0000-0003-0533-8863
June Hong Ahn
https://orcid.org/0000-0001-7104-8325
Sung-Ryeol Kim
https://orcid.org/0000-0001-7418-0049
Hyun Jung Kim
https://orcid.org/0000-0002-1878-1111

ABSTRACT

Background: Global Lung Function Initiative (GLI)-2012 reference equation is currently suggested for interpretation of spirometry results and a new local reference equation has been developed in South Korea. However, lung function profiles according to the different reference equations and their clinical relevance have not been identified in chronic obstructive pulmonary disease (COPD) patients.

Methods: Our cross-sectional study evaluated Choi's, Korean National Health and National Examination Survey (KNHANES)-VI, and GLI-2012 reference equations. We estimated the percentages of predictive forced expiratory volume in one second (FEV_1) and airflow limitation severity according to reference equations and analyzed their associations with patient reported outcomes (PROs): COPD assessment test (CAT) score, St. George's Respiratory Questionnaire for COPD patients (SGRQ-C) score, and six minute walk distance (6MWD).

Results: In the eligible 2,180 COPD patients, lower predicted values of FEV_1 and forced vital capacity (FVC) were found in GLI-2012 compared to Choi's and KNHANES-VI equations. GLI-2012 equation resulted in a lower proportion of patients being classified as FEV_1 < 80% or FVC < 80% compared to the other equations. However, the Z-scores of FEV_1 and FVC were similar between the KNHANES-VI and GLI-2012 equations. Three reference equations exhibited significant associations between FEV_1 (%) and patient-reported outcomes (CAT score, SGRQ-C score, and 6MWD).

Conclusion: GLI-2012 reference equation may not accurately reflect FEV_1 (%) in the Korean population, but the Z-score using GLI-2012 equation can be a viable option for assessing FEV_1 and airflow limitation in COPD patients. Similar to the other two equations, the GLI-2012 equation demonstrated significant associations with PROs.

Ki-Suck Jung
https://orcid.org/0000-0002-6878-6543
Kwang Ha Yoo
https://orcid.org/0000-0001-9969-2657
Deog Kyeom Kim
https://orcid.org/0000-0001-9379-8098

Funding
This work was supported by the Research Program funded Korea National Institute of Health (Fund CODE 2016ER670100, 2016ER670101, 2016ER670102, 2018ER67100, 2018ER67101, 2018ER67102, and 2021ER120500).

Disclosure
The authors have no potential conflicts of interest to disclose.

Data Availability Statement
The protocol, consent form, statistical analysis plan, definition and derivation of clinical characteristics and outcomes, training materials, regulatory documents, and other relevant study materials are available. The Institutional Review Boards will review the plan for data usage and facilitate the use of the study data. Any proposed research plan needs to satisfy high quality, follow the consent documentation and ethical approvals, and be compliant with relevant legal and regulatory requirements. The approval for data sharing will not be unreasonably withheld. Deidentified participant data will be made available within 6 months of publication.

Author Contributions
Conceptualization: Lee HW, Kim DK. Data curation: Lee HW, Lee JK, Hwang YI, Seo H, Ahn JH, Kim SR, Kim HJ, Jung KS, Yoo KH, Kim DK. Formal analysis: Lee HW, Lee JK, Hwang YI, Kim DK. Supervision: Kim DK. Writing - original draft: Lee HW. Writing - review & editing: Lee JK, Hwang YI, Seo HW, Ahn JH, Kim SR, Kim HJ, Jung KS, Yoo KH, Kim DK.

Keywords: Spirometry; Pulmonary Disease, Chronic Obstructive; Reference Values; Forced Expiratory Volume; Patient Reported Outcome Measures

INTRODUCTION

Spirometric assessment is crucial for diagnosing and assessing the severity of airflow obstruction in the patients with chronic obstructive pulmonary disease (COPD).[1] To interpret spirometry results accurately, a reference equation is necessary, as lung function can vary based on factors such as age, sex, height, thoracic wall size, and ethnicity. Recent 2022 European Respiratory Society (ERS)/American Thoracic Society (ATS) statement recommends a composite Global Lung Initiative (GLI)-2012 reference equation regardless of race/ethnicity.[2] In the US population, GLI-2012 reference equation was not inferior in predicting lung function compared to National Health and Nutrition Examination Survey III reference equation.[3] In European population, GLI-2012 reference equation was better suited than ECSC reference equation.[4] In South Korea, GLI-2012 reference equation showed a comparable performance compared with race-specific reference equations in general population.[5]

The difference in lung function according to race or ethnicity is not due to biologic difference but may reflect socioeconomic status and represent health disparities.[6] During recent decades, anthropometric and socioeconomic status has been dramatically changed in South Korea along with high level of economic growth.[7-9] Choi's reference equation, which was developed based on a database of the Korean population in 2005, has recently been reported to overestimate forced expiratory volume in one second (FEV_1) and forced vital capacity (FVC).[10,11] Recently, a new spirometric reference equation was developed based on the current database of the Korean National Health and National Examination Survey (KNHANES)-VI. The KNHANES-VI reference equation showed a higher accuracy for the predictive values of FEV_1 compared with Choi's reference equation.[5] However, it remains uncertain which reference equation for lung function is more appropriate in identifying airflow limitation in Korean COPD patients. In addition, we have insufficient evidence on the correlationship between each reference equation and patient reported outcomes (PROs) in COPD patients.

Our study aims to examine the lung function patterns according to different reference equations and to find out their correlations with PROs in COPD patients.

METHODS

We followed the statement of Strengthening the Reporting of Observational Studies in Epidemiology.[12]

Study design and eligibility criteria

This cross-sectional study analyzed the COPD patients who were registered in the Korea COPD Subgroup Study (KOCOSS) cohort from January 2012 to December 2019. The KOCOSS cohort (NCT02800499) is a prospective study including adults ≥ 40 years old with post-bronchodilator FEV_1/FVC ratio < 0.7 at 54 medical centers in South Korea. The detailed information on the KOCOSS cohort was previously reported.[13] The eligibility criteria were the patients who 1) underwent anthropometric assessments including age, sex, weight, and height, and 2) performed spirometric evaluation at the baseline visit.

Variables

All baseline information was collected upon patients' registration in the KOCOSS cohort. Sociodemographic information included age, sex, body mass index, area of residence, years of education, smoking status, Charlson comorbidity index, and history of previous lung diseases such as asthma and tuberculosis. We checked symptoms and quality of life including the COPD assessment test (CAT) score, St. George's Respiratory Questionnaire for COPD patients (SGRQ-C) score, and the 6-minute walking distance (6MWD). We identified previous history of total and severe exacerbations.

Baseline spirometric examination was performed including post-bronchodilator FEV_1, FVC, and FEV_1/FVC. The percentages of predictive FEV_1 and FVC were estimated according to three different reference equations: Choi's, KNHANES-VI, and GLI-2012 for northeast Asian. The severity of airflow limitation was classified as mild, moderate, severe, and very severe based on Global Initiative for Chronic Obstructive Lung Disease (GOLD) severity and Z-score.[14] Z-score of −1.65 was defined as lower limit of normal (LLN).

In the patients who were followed up for 3 years, CAT score, SGRQ-C score, and FEV_1 were annually recorded. We estimated an annual change of CAT score and SGRQ-C score and an annualized percentage change (mL/yr and %/yr) from the baseline FEV_1 in each individual.[15] In addition, we monitored acute exacerbations and mortality for 3 years.

Outcomes

Our primary outcome was the assessment of the percentage of predicted value of FEV_1 and the severity of airflow limitation, which was estimated using three reference equations. As secondary outcome, we evaluated the association between FEV_1 and the PROs including CAT score, SGRQ-C score, and 6MWD.

Statistical analyses

We performed Student's *t*-test or the Wilcoxon rank-sum test to evaluate continuous variables. Linear regression analyses were conducted to find the association between continuous variables. Agreement rate, Cohen's kappa, and quadratic weighted kappa were evaluated between the different reference equations for FEV_1. Given that the severity group is an ordinal variable, we prioritized quadratic weighted kappa over agreement rate or Cohen's kappa. For interpretation, we followed the criteria described in a prior paper.[16] Adding the Z-score evaluation to the predicted percentage of FEV_1 can provide objective insights into assessing an individual's relative lung function compared to the reference population, while considering age and sex.[17] For calculation of Z-score or LLN of FEV_1, we used reference data for Choi's and KNHANES-VI reference equations and R package "rspiro" for GLI-2012 reference equation. $P < 0.050$ was considered as statistical significance. All statistical analyses were performed using R statistical software, version 4.1.0 (R Core Team [2020], Vienna, Austria).

Ethics statement

The study protocol was approved by the Institutional Review Board (IRB) at each hospital (Seoul Metropolitan Government-Seoul National University Boramae Medical Center IRB No. 06-2012-36). All participants submitted their written informed consent at study enrolment. Our study was performed in accordance with the principles of the Declaration of Helsinki.

RESULTS

Baseline characteristics of the eligible 2,180 patients with COPD were described in **Table 1**. Their mean age was 69.0 and 92.6% were male patients and 92% were ever-smokers. In respiratory comorbidities, 28.9% had asthma history and 24.5% had tuberculosis history. Their median CAT score was 13 and median SGRQ-C score was 26.9. Moderate-to-severe exacerbation was found in 19.3%.

Spirometric profiles

Spirometric profiles were summarized according to the different reference equations in **Table 2**. Mean post-bronchodilator FEV_1 was 1.69 L. Mean percentage of predicted value of FEV_1 was variable according to the different reference equations: 58.5% in Choi's, 56.2% in KNHANES-VI, and 79.0% in GLI-2012 reference equations. The proportions of predictive value of FEV_1 < 80% were higher in Choi's and KNHANES-VI compared with GLI-2012 reference equations. A Z-score of FEV_1 < LLN was found in 82.1% analyzed by Choi's, 94.4% analyzed by KNHANES-VI and 92.3% analyzed by GLI-2012 reference equations. Based on Z-score of FEV_1, severe airflow limitation was more found in Choi's while moderate airflow limitation was more found in KNHANES-VI and GLI-2012 reference equations.

Mean post-bronchodilator FVC was 3.22L. The proportion of predictive value of FVC < 80% was found in 46.3% analyzed by Choi's, in 39.5% analyzed by KNHANES-VI, and in 7.2% analyzed by GLI-2012 reference equations. A Z-score of FVC < LLN was found in 44.3%

Table 1. Baseline characteristics of the included COPD patients

Characteristics	COPD patients (N = 2,180)
Age	69.0 ± 7.7
< 65	598 (27.4)
65–69	491 (22.5)
≥ 70	1,091 (50.0)
Male	2,019 (92.6)
Body mass index	23.0 ± 3.4
Residence	
Rural area	792 (36.3)
School years	
≤ 6 years	702 (32.2)
7–12 years	287 (13.2)
> 12 years	1,191 (54.6)
Smoking status	
Never smoker	175 (8.0)
Ex-smoker	1,417 (65.0)
Current smoker	588 (27.0)
Charlson comorbidity index	1 (1, 1)
0–1	1,643 (75.4)
2–3	430 (19.7)
≥ 4	20 (0.9)
Asthma history	630 (28.9)
Tuberculosis history	534 (24.5)
CAT score	13 (8, 20)
SGRQ-C score	26.9 (16.0, 44.7)
6MWD	390 (319, 458)
Previous exacerbation	
Moderate-to-severe exacerbation	420 (19.3)
Severe exacerbation	206 (9.4)

Values are presented as mean ± standard deviation or median (interquartile range) or number (%).
COPD = chronic obstructive pulmonary disease, CAT = COPD assessment test, SGRQ-C = St. George's Respiratory Questionnaire for COPD patients, 6MWD = 6-minute walking distance.

Table 2. Spirometric profiles according to the different reference equations

Characteristics	Absolute value	Reference equations		
		Choi's	KNHANES-VI	GLI-2012
Post-BDR FEV_1, L	1.69 ± 0.59			
% of predicted FEV_1		58.5 ± 18.2	56.2 ± 18.5	79.0 ± 24.8
GOLD severity[a]				
Stage I		264 (12.1)	297 (13.6)	1,049 (48.1)
Stage II		1,186 (54.4)	1,159 (53.2)	832 (38.2)
Stage III		615 (28.2)	618 (28.3)	280 (12.8)
Stage IV		115 (5.3)	106 (4.9)	19 (0.9)
Z-score of FEV_1		−3.50 (−4.91, −2.18)	−3.31 (−4.0, −2.6)	−3.07 (−4.27, −1.91)
< LLN; 5th percentile		1,790 (82.1)	2,059 (94.4)	2,013 (92.3)
Mild (Z-score: −1.65, −2.5)		286 (13.1)	345 (15.8)	319 (14.6)
Moderate (Z-score: −2.51, −4.0)		618 (28.3)	1,174 (53.9)	1,024 (47.0)
Severe (Z-score: ≤ −4.1)		886 (40.6)	540 (24.8)	670 (30.7)
Post-BDR FVC, L	3.22 ± 0.82			
% of predicted FVC		81.2 ± 16.6	84.6 ± 17.5	117.1 ± 24.2
% of predicted FVC < 80%		1,006 (46.3)	862 (39.5)	156 (7.2)
Z-score of FVC		−1.39 (−2.64, −0.10)	−0.73 (−1.35, −0.06)	−0.20 (−1.50, 1.06)
< LLN; 5th percentile		965 (44.3)	367 (16.8)	480 (22.0)
Post-BDR FEV_1/FVC, %	50.9 ± 11.9			
< LLN; 5th percentile		2,032 (93.2)	2,163 (99.2)	2,045 (93.8)
BDR[b]	406 (18.6)			
DL_{CO}, %	63.9 ± 20.7			
DL_{CO}/VA, %	75.6 ± 23.9			
TLC, L	5.62 ± 0.85			
TLC, %	95.0 ± 13.2			

Values are presented as mean ± standard deviation or median (interquartile range) or number (%).
KNHANES = Korean National Health and Nutritional Examination Survey, GLI = Global Lung Function Initiative, BDR = bronchodilator reversibility, FEV_1 = forced expiratory volume in one second, GOLD = Global Initiative for Chronic Obstructive Lung Disease, LLN = lower limit of normal, FVC = forced vital capacity, DL_{CO} = diffusing capacity for carbon monoxide, VA = alveolar volume, TLC = total lung capacity.
[a]Stage I = predicted FEV_1 ≥ 80%, stage II = predicted FEV_1 < 80% & predicted FEV_1 ≥ 50%, stage III = predicted FEV_1 < 50% & predicted FEV_1 ≥ 30%, stage IV = predicted FEV_1< 30%.
[b]FEV_1 improvement from the pre-bronchodilator value by 12% and >200 mL.

analyzed by Choi's, in 16.8% analyzed by KNHANES-VI, in 22.0% analyzed by GLI-2012 reference equations.

Mean post-bronchodilator FEV_1/FVC was 50.9%. A Z-score of FEV_1/FVC < LLN was found in 93.2% analyzed by Choi's, 99.2% analyzed by KNHANES-VI, and 93.8% analyzed by GLI-2012 reference equations.

Concordance of airflow limitation severity

There were strong linear relationships between different reference equations (**Supplementary Fig. 1**). In the analysis for the severity of airflow limitation based on GOLD classification, strong agreement was found between Choi's and KNHANES-VI reference equations (agreement rate = 92.7%, Cohen's kappa = 0.88, weighted kappa = 0.93), while weak agreement was found between GLI-2012 and Choi's or GLI-2012 and KNHANES-VI reference equations (**Table 3**, **Supplementary Tables 1** and **2**). In analysis for the severity of airflow limitation based on Z-score, moderate agreement was found among the three reference equations.

PROs and different reference equations

There was a negative linear relationship between CAT score and the percentage of predicted value of FEV_1 estimated by the three reference equations (**Fig. 1**). CAT score was significantly different according to the severity of airflow limitation based on GOLD classification estimated by the three reference equations. There was a negative linear relationship between

Table 3. Concordance of airflow limitation severity among different reference equations of FEV_1

Reference equations for FEV_1	Agreement rate	Cohen's kappa	Weighted kappa
Severity of airflow limitation based on GOLD classification			
Choi's–KNHANES-VI	92.7%	0.88 (0.86–0.90)	0.93 (0.93–0.93)
Choi's–GLI-2012	39.8%	0.14 (0.11–0.16)	0.58 (0.52–0.64)
KNHANES-VI–GLI-2012	42.4%	0.17 (0.14–0.20)	0.59 (0.53–0.65)
Severity of airflow limitation based on Z-score			
Choi's–KNHANES-VI	51.4%	0.32 (0.29–0.35)	0.64 (0.63–0.65)
Choi's–GLI-2012	59.5%	0.43 (0.40–0.46)	0.64 (0.62–0.65)
KNHANES-VI–GLI-2012	60.7%	0.39 (0.36–0.42)	0.69 (0.69–0.69)

FEV_1 = forced expiratory volume in one second, GOLD = Global Initiative for Chronic Obstructive Lung Disease, KNHANES = Korean National Health and Nutritional Examination Survey, GLI = Global Lung Function Initiative.

CAT score and Z-score of FEV_1 estimated by the three reference equations (**Supplementary Fig. 2**). Differences in CAT score according to mild, moderate, and severe airflow limitation based on Z-score of FEV_1 were well discriminated in KNAHNES VI reference equation.

There was a negative linear relationship between SGRQ-C score and the percentage of predicted value of FEV_1 estimated by the three reference equations (**Fig. 2**). SGRQ-C score was significantly different according to the severity of airflow limitation based on GOLD classification estimated by the three reference equations. There was a negative linear relationship between SGRQ-C score and Z-score of FEV_1 estimated by different reference equations (**Supplementary Fig. 3**). Differences in SGRQ-C scores according to the severity of airflow limitation based on Z-score of FEV_1 were well discriminated in Choi's reference equation.

There was a positive linear relationship between 6MWD and the percentage of predicted value of FEV_1 estimated by the three reference equations (**Fig. 3**). 6MWD was significantly different between GOLD stage I and II only when estimated by GLI-2012 equation while between GOLD stage III and IV when estimated by Choi's and KNAHNES VI reference equations. There was a positive linear relationship between 6MWD and Z-score of FEV_1 estimated by different reference equations (**Supplementary Fig. 4**). Differences in 6MWD according to mild, moderate, and severe airflow limitation based on Z-score of FEV_1 were well discriminated in KNAHNES VI reference equation.

DISCUSSION

Our study compared the three reference equations (Choi's, KNHANES-VI, and GLI-2012) in COPD patients. The GLI-2012 equation yielded lower predicted values of FEV_1 and FVC compared to Choi's and KNHANES-VI equations. It also resulted in a lower proportion of patients being classified as FEV_1 < 80% or FVC < 80% compared to the other equations. However, the Z-scores of FEV_1 and FVC were similar between the KNHANES-VI and GLI-2012 equations. The GLI-2012 equation exhibited significant associations between FEV_1 (%) or the severity of airflow limitation and patient-reported outcomes (CAT score, SGRQ-C score, and 6MWD). These findings suggest that the Z-score using GLI-2012 equation can be a viable option for assessing lung function in COPD patients in South Korea.

In spirometric interpretation, fixed cut-offs often overlook a significant limitation, namely that the normal range can vary depending on age, gender, and population characteristics.[18,19] The GOLD report has conventionally determined airflow limitation as the percentage of predicted value of FEV1 <80% and airway obstruction as fixed ratio of FEV_1/FVC < 0.7.[14] This

https://doi.org/10.3346/jkms.2024.39.e20

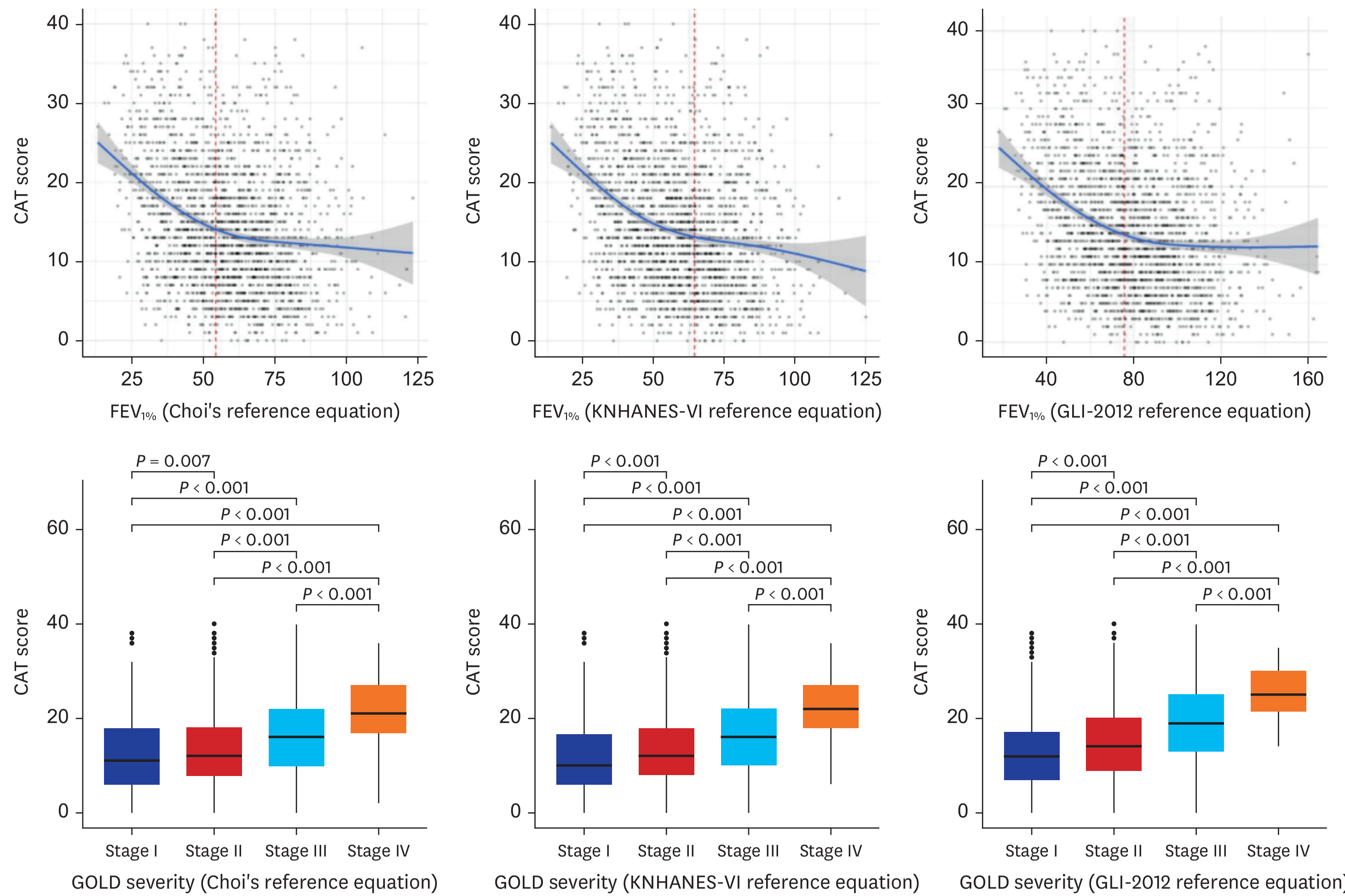

Fig. 1. Association between CAT score and different reference equations of FEV$_1$. The vertical dotted line indicates the optimal cut-off value of FEV$_{1\%}$ to predict CAT ≥ 10. R^2 was 0.078 (P < 0.001) between CAT score and FEV$_{1\%}$ (Choi's reference equation), and 0.084 (P < 0.001) between CAT score and FEV$_{1\%}$ (KNHANES-VI reference equation), and 0.073 (P < 0.001) between CAT score and FEV$_{1\%}$ (GLI-2012 reference equation).

CAT = COPD assessment test, FEV$_1$ = forced expiratory volume in one second, GOLD = Global Initiative for Chronic Obstructive Lung Disease, KNHANES = Korean National Health and Nutritional Examination Survey, GLI = Global Lung Function Initiative.

JKMS

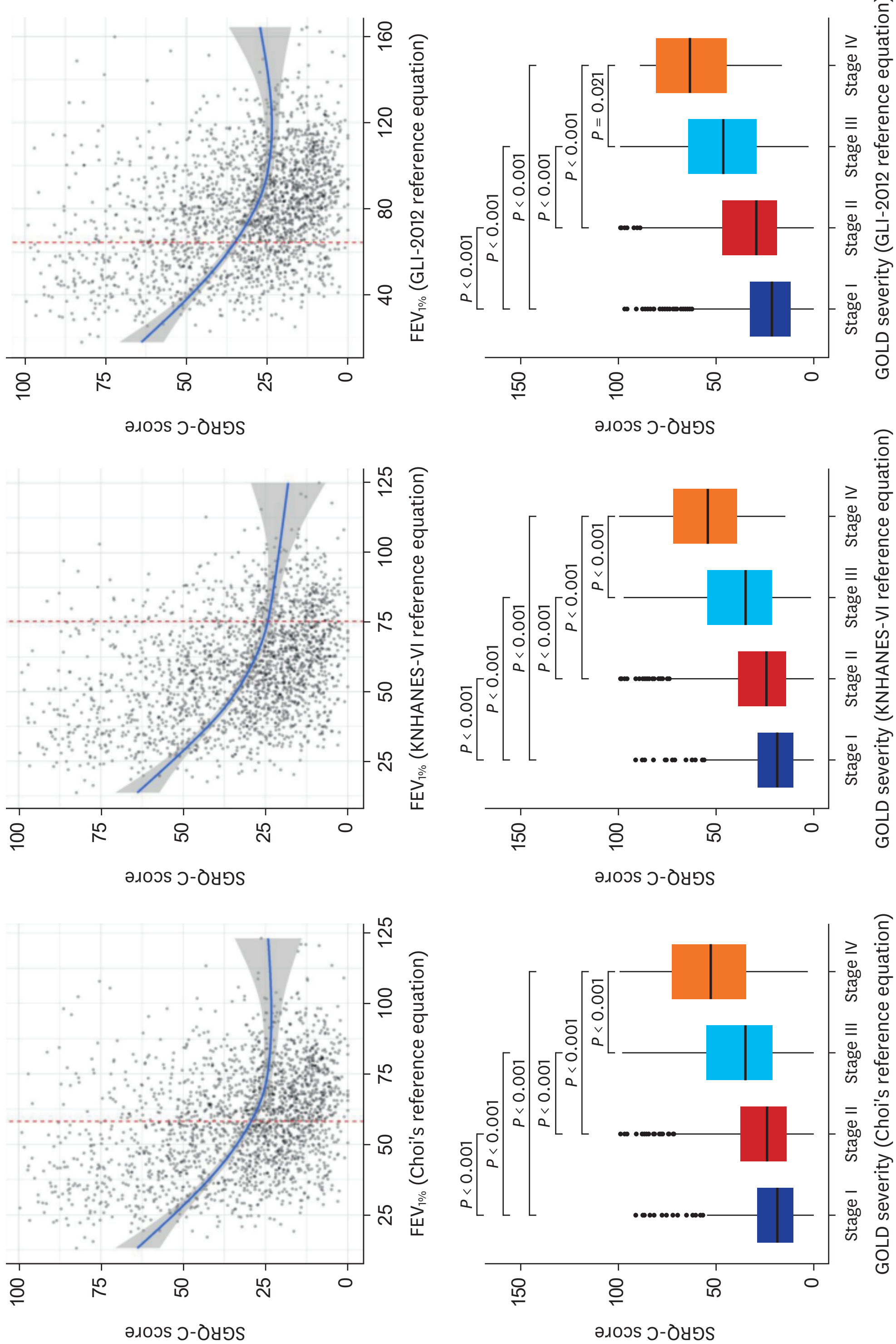

Fig. 2. Association between SGRQ-C score and different reference equations of FEV_1. The vertical dotted line indicates the optimal cut-off value of $FEV_{1\%}$ to predict SGRQ-C score ≥20. R^2 was 0.132 (P < 0.001) between SGRQ-C score and $FEV_{1\%}$ (Choi's reference equation), and 0.145 (P < 0.001) between SGRQ-C score and $FEV_{1\%}$ (KNHANES-VI reference equation), and 0.128 (P < 0.001) between SGRQ-C score and $FEV_{1\%}$ (GLI-2012 reference equation).

SGRQ-C = St. George's Respiratory Questionnaire for COPD patients, FEV_1 = forced expiratory volume in one second, GOLD = Global Initiative for Chronic Obstructive Lung Disease, KNHANES = Korean National Health and Nutritional Examination Survey, GLI = Global Lung Function Initiative.

 https://doi.org/10.3346/jkms.2024.39.e20

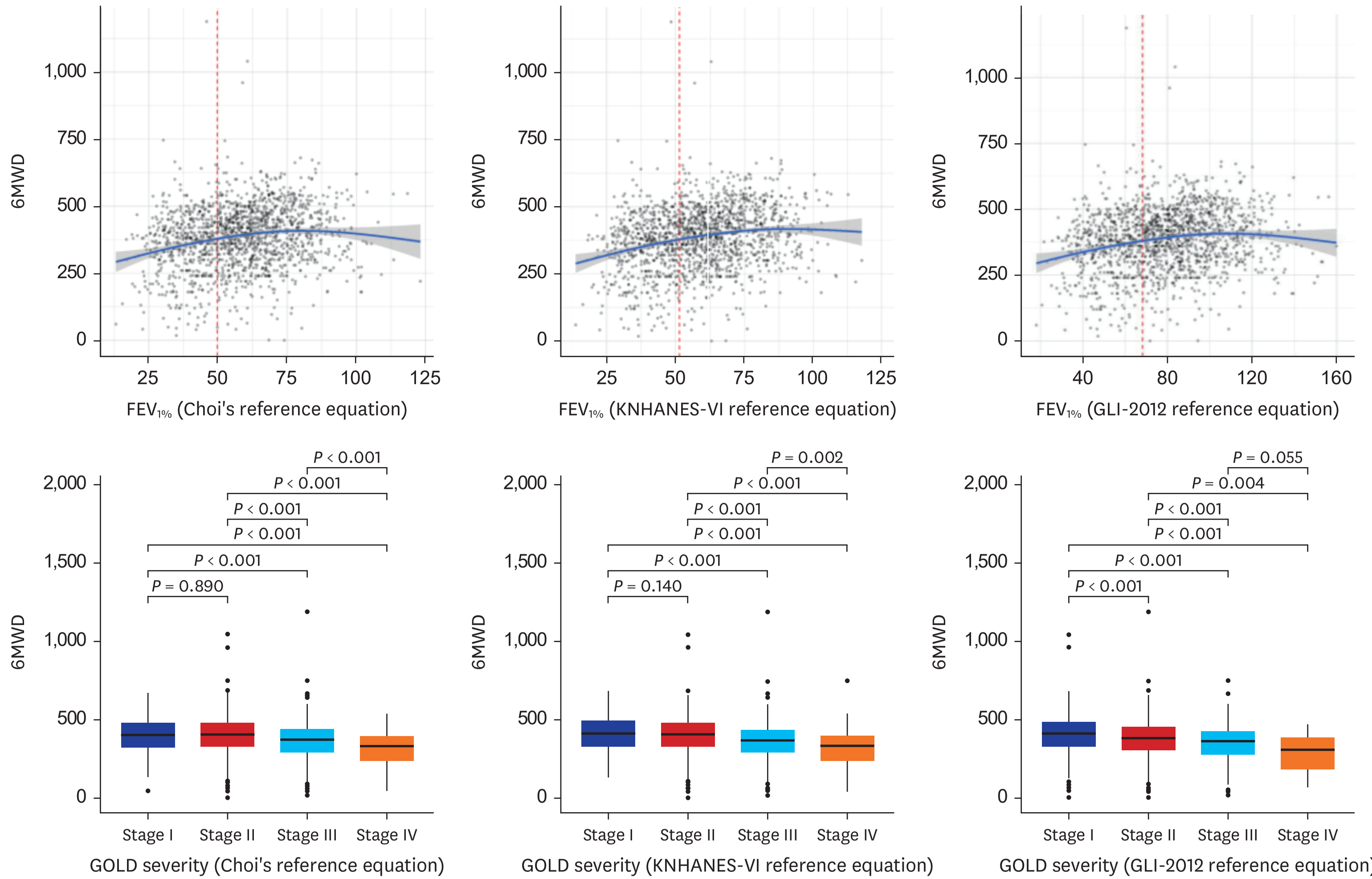

Fig. 3. Association between 6MWD and different reference equations of FEV_1. The vertical dotted line indicates the optimal cut-off value of $FEV_{1\%}$ to predict 6MWD ≥ 400. R^2 was 0.033 ($P < 0.001$) between 6MWD and $FEV_{1\%}$ (Choi's reference equation), and 0.035 ($P < 0.001$) between 6MWD and $FEV_{1\%}$ (KNHANES-VI reference equation), and 0.050 ($P < 0.001$) between 6MWD and $FEV_{1\%}$ (GLI-2012 reference equation).
6MWD = 6-minute walking distance, FEV_1 = forced expiratory volume in one second, GOLD = Global Initiative for Chronic Obstructive Lung Disease, KNHANES = Korean National Health and Nutritional Examination Survey, GLI = Global Lung Function Initiative.

approach has been favored for its simplicity in calculation and ease of data collection and interpretation. However, the cut-off based on the percentage of predicted value of FEV_1 or the ratio of FEV_1/FVC is derived from the average values of a general population and does not consider that the normal range can vary with age, sex, and other demographic characteristics. A previous study revealed that in older individuals with smaller body sizes, the percentage of predicted FEV_1 can be overestimated even when domestic reference equation was used.[11] Additionally, it should be noted that the percentage of predicted value of FEV_1 and FEV_1/FVC ratio naturally decreases by 1.24% point and 0.32% point, respectively, with age in the general population.[20,21] This suggests that if airway obstruction or airflow limitation is determined based on a conventional fixed value, it can be underestimated in younger individuals and overestimated in older individuals.

The preference for Z-scores or LLN arises from their distinctive capacity to provide a more precise determination of whether an individual's lung function is within the normal range, in comparison to the results of healthy individuals.[2] The Z-scores or the LLN enable the estimation of the probability that an individual's lung function falls outside the normal range, taking into account their characteristics such as age, gender, height, and ethnicity.[2] In spirometric assessments, the 5th percentile of Z-scores is often used as the LLN cut-off, indicating a 5% probability that an individual's lung function falls below the normal range. Our study observed a significant disparity between GLI-2012 and Choi's and KNHANES-VI, when we defined airflow limitation or airway obstruction using fixed cut-off values. However, this discrepancy was substantially reduced when the 5th percentile of Z-scores was set for the LLN. This finding is aligned with a previous study that reported a high agreement in airflow limitation severity based on Z-scores while low agreement based on the percentage of predicted value of FEV_1.[22]

Currently, GLI reference equation has been recommended for spirometric evaluation regardless of geographic or ethnic background. It seems a meaningful step to evaluate the lung function of the world's population with a universal reference equation. Through unified reference equation, we can generalize the results from clinical research in different countries and present comprehensive practical guidance. However, there are caveats to the application of the GLI-2012 reference equation to the Korean population. In our study, we found that the Z-scores of the GLI-2012 reference equation closely aligned with those of the KNHANES-VI reference equation, but it tended to underestimate the abnormality of FEV_1 and $FEV_{1/}$ FVC. Similar observations were made in another study involving 1,243 Italian children with normal lung function. While the exact reasons for these discrepancies remain unclear, we hypothesize two possible factors.[23] Firstly, in the Korean birth cohort, the GLI-2012 reference equation showed a significant discrepancy with real-world values.[24] This discrepancy is attributed to the exclusion of individuals under the age of 15 in the northeast Asian population during its development.[25] Secondly, the GLI-2012 reference equation for the northeast Asian was developed by amalgamating the populations of North China and South Korea without differentiation.[25] Consequently, it is challenging to assert that the GLI-2012 reference equation is optimally tailored to Koreans.

In general, predictive value of lung function is expected to be more accurate when using a reference equation developed from a population with similar demographic characteristics.[3] A recent study suggested higher accuracy when the information on region was added to the lung function reference equation.[26] In fact, our study showed considerable discrepancy in the predictive value of FEV_1 or COPD severity between global reference equation (GLI-2012) and region-specific reference equations (Choi's and KHANES VI). One of the potential reasons for the discrepancy may be attributed to variations in human body proportions based on different ethnicities of GLI-2012 reference equation. A longer upper body segment length seems to be associated with larger thoracic cage size among different ethnicities.[27] Given that Asians have a higher ratio of sitting height to standing height compared to Caucasians, thoracic volume may actually be larger in Asians than in Caucasians of the same height.[28,29] As a result, the GLI-2012 equation might underestimate the predicted values of FEV_1 or FVC in Asians compared to their actual values. Another potential reason for the discrepancy is the differences in the methods used to develop the reference equations. The GLI-2012 reference equation was created using the lambda, mu, sigma method with the specific goal of predicting Z-scores, whereas Choi's and KNHANES-VI reference equations were developed for predicting values of lung function parameters using conventional regression analysis.

Choi's reference equation, developed in 2005, is commonly used in South Korea.[10] However, the spirometric reference equation developed from the KNHANES-VI population differed from Choi's reference equation.[5] KNHANES-VI reference equation added more weights on age compared to Choi's reference equation. The predictive value of lung function using the KNHANES-VI reference equation closely matched the actual lung function, whereas Choi's reference equation exhibited a significant margin of error.[5] This suggests that the reference equation for lung function needs to be updated considering the demographic variations across different time periods. In our study, spirometric airway obstruction, as determined by Z-score (post-bronchodilator reversibility [BDR] FEV_1/FVC < LLN) was more sensitively identified using KNHANES-VI reference equation compared to Choi's reference equation. Therefore, if LLN is used to diagnose COPD, spirometric evaluation using the KNHANES-VI reference equation may lead to the identification of more patients with spirometric airway obstruction. On the other hand, the GLI-2012 reference equation appears to be less sensitive in detecting spirometric airway obstruction (post-BDR FEV_1/FVC < LLN) in Koreans. As this finding raises concerns about the underdiagnosis of COPD, additional adjustment may be necessary for the practical application of the GLI-2012 reference equation.

When applying the KNHANES-VI reference equation to COPD patients, it is essential to be aware of several limitations. The KNHANES-VI reference equation was developed based on a sample of 117 individuals, which is relatively small to claim complete representation of the general population in Korea. Also, there is a noticeable disparity in gender and age distribution between the population used to develop the KNHANES VI prediction equation, where a higher proportion of females and a lower proportion of elderly individuals were included, and the demographics commonly observed among COPD patients, who are typically predominantly male and older.

In COPD patients, since FEV_1 alone cannot fully represent the functional impairment and respiratory symptoms, the evaluation of PROs is crucial for a comprehensive assessment.[30,31] FEV_1 and PROs were correlated in the patients with COPD. Negative linear relationships have been reported between CAT score and FEV_1 (%) ($r = -0.34, -0.55$).[32,33] CAT scores were higher in COPD patients with a higher severity of airflow limitation.[33] The association between CAT score and FEV_1 (%) was more prominent in symptomatic COPD patients or those during acute exacerbation of COPD.[34,35] Rapid FEV_1 decline was associated with severe dyspnea in COPD patients.[36] However, when evaluating treatment response, there was no correlation between FEV_1 and CAT.[37] SGRQ was negatively correlated with Post-bronchodilator FEV_1 (%).[38] The absolute difference of SGRQ score was about 13 between GOLD 1/2 and 3/4.[39] In a meta-analysis, FEV_1 (mL) and SGRQ score showed a significant correlationship and their pooled correlation coefficient was −0.63.[40] The association between 6MWD and FEV_1 has not been clearly identified. In a study, 6MWD was associated with FEV_1 ($r = 0.260$),[41] but another study showed no significant relationship between 6MWD and FEV_1.[42] In our study, CAT score, SGRQ-C score, and 6MWD were correlated with predictive percentage of FEV_1 (%) or GOLD severity, regardless of the different reference equations in COPD patients. In the analyses for CAT and SGRQ-C scores, more linear relationship was found with Z-score of FEV_1 while more pronounced difference was found according to COPD severity based on the predictive percentage of FEV_1 (%).

Our study has several limitations. First, the accuracy, especially the specificity, of airflow limitation severity cannot be confirmed since a healthy population was not used as a control. Second, it remains unclear whether the difference between each reference equation has an impact on clinical practice. The definition of COPD has mainly been based on the FEV_1/FVC ratio, and current clinical practice does not differ according to the airflow limitation severity. However, considering that FEV_1 (%) is an important indicator for evaluating clinical prognosis, using a more accurate reference equation would be more beneficial in assessing COPD patients.[43] Third, we used the GOLD criteria to evaluate airflow limitation severity which was developed only using GLI-2012 reference equation. In GOLD severity, the GLI-2012 reference equation shows a significant difference from Choi's or KNHANES-VI. On the other hand, in the Z-score severity, the difference between the three reference equations was greatly reduced. As recent 2022 ERS/ATS statement suggested, it seems appropriate to use Z-score when globally evaluating lung function with GLI-2012 reference equation. Fourth, it is difficult to generalize that PROs and FEV_1 are correlated. In our study, the linear association between PROs and FEV_1 (%) seemed clearer until a certain cut-off, and then the association seemed weaker.

In conclusion, our findings highlight the need for careful consideration in selecting an appropriate reference equation for spirometry interpretation in COPD patients. GLI-2012 reference equation may not accurately reflect FEV_1 (%) in the Korean population, but the Z-score using GLI-2012 equation can be a viable option for assessing FEV_1 and airflow limitation in COPD patients. KNHANES-VI reference equation seemed to be more efficient to detect undiagnosed COPD using the LLN. Similar to the other two equations specific for Korean, the GLI-2012 equation demonstrated significant associations with PROs.

ACKNOWLEDGMENTS

Data in this study were from the Korea COPD Subgroup Study (KOCOSS) cohort (NCT02800499), a prospective study including adults ≥ 40 years old with post-bronchodilator forced expiratory volume in one second (FEV_1)/forced vital capacity (FVC) ratio < 0.7 at 54 medical centers in South Korea.

SUPPLEMENTARY MATERIALS

Supplementary Table 1

Severity of airflow limitation based on GOLD report in different reference equations

Supplementary Table 2

Severity of airflow limitation based on Z-score in different references

 https://doi.org/10.3346/jkms.2024.39.e20

Supplementary Fig. 1

Association among different reference equations of FEV_1.

Supplementary Fig. 2

Association between CAT score and Z-score estimated by different reference equations of FEV_1. The vertical dotted line indicates the optimal cut-off value of Z-score of $FEV_{1\%}$ to predict CAT ≥ 10. R^2 was 0.066 (P < 0.001) between CAT score and Z-score of $FEV_{1\%}$ (Choi's reference equation), and 0.078 (P < 0.001) between CAT score and Z-score of $FEV_{1\%}$ (KNHANES-VI reference equation), and 0.072 (P < 0.001) between CAT score and Z-score of $FEV_{1\%}$ (GLI-2012 reference equation).

Supplementary Fig. 3

Association between SGRQ-C score and Z-score estimated by different reference equations of FEV_1. The vertical dotted line indicates the optimal cut-off value of Z-score of $FEV_{1\%}$ to predict SGRQ-C ≥20. R^2 was 0.104 (P < 0.001) between SGRQ-C score and Z-score of $FEV_{1\%}$ (Choi's reference equation), and 0.144 (P < 0.001) between SGRQ-C score and Z-score of $FEV_{1\%}$ (KNHANES-VI reference equation), and 0.120 (P < 0.001) between SGRQ-C score and Z-score of $FEV_{1\%}$ (GLI-2012 reference equation).

Supplementary Fig. 4

Association between 6MWD and Z-score estimated by different reference equations of FEV_1. The vertical dotted line indicates the optimal cut-off value of Z-score of $FEV_{1\%}$ to predict 6MWD ≥ 400. R^2 was 0.013 (P < 0.001) between 6MWD and Z-score of $FEV_{1\%}$ (Choi's reference equation), and 0.077 (P < 0.001) between 6MWD and Z-score of $FEV_{1\%}$ (KNHANES-VI reference equation), and 0.022 (P < 0.001) between 6MWD and Z-score of $FEV_{1\%}$ (GLI-2012 reference equation).

REFERENCES

1. Rennard S, Thomashow B, Crapo J, Yawn B, McIvor A, Cerreta S, et al. Introducing the COPD foundation guide for diagnosis and management of COPD, recommendations of the COPD foundation. *COPD* 2013;10(3):378-89. PUBMED | CROSSREF
2. Stanojevic S, Kaminsky DA, Miller MR, Thompson B, Aliverti A, Barjaktarevic I, et al. ERS/ATS technical standard on interpretive strategies for routine lung function tests. *Eur Respir J* 2022;60(1):2101499. PUBMED | CROSSREF
3. Haynes JM, Kaminsky DA, Stanojevic S, Ruppel GL. Pulmonary function reference equations: a brief history to explain all the confusion. *Respir Care* 2020;65(7):1030-8. PUBMED | CROSSREF
4. Hulo S, de Broucker V, Giovannelli J, Cherot-Kornobis N, Nève V, Sobaszek A, et al. Global Lung Function Initiative reference equations better describe a middle-aged, healthy French population than the European Community for Steel and Coal values. *Eur Respir J* 2016;48(6):1779-81. PUBMED | CROSSREF
5. Choi HS, Park YB, Yoon HK, Lim SY, Kim TH, Park JH, et al. Validation of previous spirometric reference equations and new equations. *J Korean Med Sci* 2019;34(47):e304. PUBMED | CROSSREF
6. Bhakta NR, Bime C, Kaminsky DA, McCormack MC, Thakur N, Stanojevic S, et al. Race and ethnicity in pulmonary function test interpretation: an official American thoracic society statement. *Am J Respir Crit Care Med* 2023;207(8):978-95. PUBMED | CROSSREF
7. NCD Risk Factor Collaboration (NCD-RisC). A century of trends in adult human height. *Elife* 2016;5:e13410. PUBMED
8. Yang YS, Han BD, Han K, Jung JH, Son JW; Taskforce Team of the Obesity Fact Sheet of the Korean Society for the Study of Obesity. Obesity fact sheet in Korea, 2021: trends in obesity prevalence and obesity-related comorbidity incidence stratified by age from 2009 to 2019. *J Obes Metab Syndr* 2022;31(2):169-77. PUBMED | CROSSREF

9. Cole TJ, Mori H. Fifty years of child height and weight in Japan and South Korea: contrasting secular trend patterns analyzed by SITAR. *Am J Hum Biol* 2018;30(1):e23054. PUBMED | CROSSREF

10. Choi JK, Paek D, Lee JO. Normal predictive values of spirometry in Korean population. *Tuberc Respir Dis (Seoul)* 2005;58(3):230-42. CROSSREF

11. Chang Y, Kim HC, Jo KW, Lee JS, Oh YM, Lee SD, et al. Analysis of high predicted pulmonary function: possibility of overestimation in small elderly examinees. *Korean J Intern Med* 2020;35(1):142-9. PUBMED | CROSSREF

12. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007;370(9596):1453-7. PUBMED | CROSSREF

13. Lee JY, Chon GR, Rhee CK, Kim DK, Yoon HK, Lee JH, et al. Characteristics of patients with chronic obstructive pulmonary disease at the first visit to a pulmonary medical center in Korea: The KOrea COpd subgroup study team cohort. *J Korean Med Sci* 2016;31(4):553-60. PUBMED | CROSSREF

14. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. http://www.goldcopd.org/. Updated 2023. Accessed October 30, 2023.

15. Mannino DM, Reichert MM, Davis KJ. Lung function decline and outcomes in an adult population. *Am J Respir Crit Care Med* 2006;173(9):985-90. PUBMED | CROSSREF

16. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22(3):276-82. PUBMED | CROSSREF

17. Stanojevic S, Wade A, Stocks J. Reference values for lung function: past, present and future. *Eur Respir J* 2010;36(1):12-9. PUBMED | CROSSREF

18. Miller MR, Quanjer PH, Swanney MP, Ruppel G, Enright PL. Interpreting lung function data using 80% predicted and fixed thresholds misclassifies more than 20% of patients. *Chest* 2011;139(1):52-9. PUBMED | CROSSREF

19. Swanney MP, Ruppel G, Enright PL, Pedersen OF, Crapo RO, Miller MR, et al. Using the lower limit of normal for the FEV_1/FVC ratio reduces the misclassification of airway obstruction. *Thorax* 2008;63(12):1046-51. PUBMED | CROSSREF

20. Choi KY, Lee HJ, Lee JK, Park TY, Heo EY, Kim DK, et al. Rapid FEV_1/FVC decline is related with incidence of obstructive lung disease and mortality in general population. *J Korean Med Sci* 2023;38(1):e4. PUBMED | CROSSREF

21. Lee HW, Lee HJ, Lee JK, Park TY, Heo EY, Kim DK. Rapid FEV_1 decline and lung cancer incidence in South Korea. *Chest* 2022;162(2):466-74. PUBMED | CROSSREF

22. Huprikar NA, Holley AB, Skabelund AJ, Hayes JA, Hiles PD, Aden JK, et al. A comparison of global lung initiative 2012 with third national health and nutrition examination survey spirometry reference values. implications in defining obstruction. *Ann Am Thorac Soc* 2019;16(2):225-30. PUBMED | CROSSREF

23. Fasola S, La Grutta S, Cibella F, Cilluffo G, Viegi G. Global Lung Function Initiative 2012 reference values for spirometry in South Italian children. *Respir Med* 2017;131:11-7. PUBMED | CROSSREF

24. Park JS, Suh DI, Choi YJ, Ahn K, Kim KW, Shin YH, et al. Pulmonary function of healthy Korean children from three independent birth cohorts: validation of the Global Lung Function Initiative 2012 equation. *Pediatr Pulmonol* 2021;56(10):3310-20. PUBMED | CROSSREF

25. 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(6):1324-43. PUBMED | CROSSREF

26. Leong WY, Gupta A, Hasan M, Mahmood S, Siddiqui S, Ahmed S, et al. Reference equations for evaluation of spirometry function tests in South Asia, and among South Asians living in other countries. *Eur Respir J* 2022;60(6):2102962. PUBMED | CROSSREF

27. Yap WS, Chan CC, Chan SP, Wang YT. Ethnic differences in anthropometry among adult Singaporean Chinese, Malays and Indians, and their effects on lung volumes. *Respir Med* 2001;95(4):297-304. PUBMED | CROSSREF

28. Nakanishi Y, Nethery V. Anthropometric comparison between Japanese and Caucasian American male university students. *Appl Human Sci* 1999;18(1):9-11. PUBMED | CROSSREF

29. Lim U, Ernst T, Buchthal SD, Latch M, Albright CL, Wilkens LR, et al. Asian women have greater abdominal and visceral adiposity than Caucasian women with similar body mass index. *Nutr Diabetes* 2011;1(5):e6. PUBMED | CROSSREF

30. Kakavas S, Kotsiou OS, Perlikos F, Mermiri M, Mavrovounis G, Gourgoulianis K, et al. Pulmonary function testing in COPD: looking beyond the curtain of FEV_1. *NPJ Prim Care Respir Med* 2021;31(1):23. PUBMED | CROSSREF

31. Lange P, Halpin DM, O'Donnell DE, MacNee W. Diagnosis, assessment, and phenotyping of COPD: beyond FEV_1. *Int J Chron Obstruct Pulmon Dis* 2016;11 Spec Iss(Spec Iss):3-12. PUBMED | CROSSREF

32. Miravitlles M, Koblizek V, Esquinas C, Milenkovic B, Barczyk A, Tkacova R, et al. Determinants of CAT (COPD assessment test) scores in a population of patients with COPD in central and Eastern Europe: the POPE study. *Respir Med* 2019;150:141-8. PUBMED | CROSSREF

33. Ghobadi H, Ahari SS, Kameli A, Lari SM. The relationship between COPD assessment test (CAT) scores and severity of airflow obstruction in stable COPD patients. *Tanaffos* 2012;11(2):22-6. PUBMED

34. Tu YH, Zhang Y, Fei GH. Utility of the CAT in the therapy assessment of COPD exacerbations in China. *BMC Pulm Med* 2014;14(1):42. PUBMED | CROSSREF

35. Huang WC, Wu MF, Chen HC, Hsu JY; TOLD Group. Features of COPD patients by comparing CAT with mMRC: a retrospective, cross-sectional study. *NPJ Prim Care Respir Med* 2015;25(1):15063. PUBMED | CROSSREF

36. Lee HW, Lee JK, Lee MG, Shin KC, Ra SW, Kim TH, et al. Risk factors of rapid FEV_1 decline in a real-world chronic obstructive pulmonary disease cohort. *Respiration* 2022;101(12):1078-87. PUBMED | CROSSREF

37. Kostikas K, Greulich T, Mackay AJ, Lossi NS, Aalamian-Mattheis M, Nunez X, et al. Treatment response in COPD: does FEV_1 say it all? A *post hoc* analysis of the CRYSTAL study. *ERJ Open Res* 2019;5(1):00243-2018. PUBMED | CROSSREF

38. Weatherall M, Marsh S, Shirtcliffe P, Williams M, Travers J, Beasley R. Quality of life measured by the St George's Respiratory Questionnaire and spirometry. *Eur Respir J* 2009;33(5):1025-30. PUBMED | CROSSREF

39. Lutter JI, Jörres RA, Kahnert K, Schwarzkopf L, Studnicka M, Karrasch S, et al. Health-related quality of life associates with change in FEV_1 in COPD: results from the COSYCONET cohort. *BMC Pulm Med* 2020;20(1):148. PUBMED | CROSSREF

40. de la Loge C, Tugaut B, Fofana F, Lambert J, Hennig M, Tschiesner U, et al. Relationship between FEV_1 and patient-reported outcomes changes: results of a meta-analysis of randomized trials in stable COPD. *Chronic Obstr Pulm Dis (Miami)* 2016;3(2):519-38. PUBMED | CROSSREF

41. Agrawal MB, Awad NT. Correlation between six minute walk test and spirometry in chronic pulmonary disease. *J Clin Diagn Res* 2015;9(8):OC01-04. PUBMED | CROSSREF

42. Chauhan NK, Sanwaria P, Singh S, Dutt N, Saini L, Kumar S, et al. Correlation between FEV_1 percentage predicted and 6 minute walk distance in patients of chronic obstructive pulmonary disease. 2017;61(1):38-42.

43. Celli BR, Cote CG, Marin JM, Casanova C, Montes de Oca M, Mendez RA, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med* 2004;350(10):1005-12. PUBMED | CROSSREF