Received: 26 October 2021 | Accepted: 7 February 2022

DOI: 10.1111/cpf.12745

ORIGINAL ARTICLE

Clinical Physiology and Functional Imaging

# Restrictive spirometry versus restrictive lung function using the GLI reference values

Tomi Myrberg[1] | Anne Lindberg[2] | Berne Eriksson[3,4] | Linnea Hedman[5] | Caroline Stridsman[2] | Bo Lundbäck[4] | Eva Rönmark[5] | Helena Backman[5]

[1]Department of Surgical and Perioperative Sciences, Anesthesiology and Intensive Care Medicine, Umeå University, Umeå, Sweden

[2]Department of Public Health and Clinical Medicine, Section of Medicine, Umeå University, Umeå, Sweden

[3]Department of Krefting Research Centre, Institute of Medicine, University of Gothenburg, Gothenburg, Sweden

[4]Department of Research and Development, Region Halland, Halmstad, Sweden

[5]Department of Public Health and Clinical Medicine, Section of Sustainable Health/the OLIN Unit, Umeå University, Umeå, Sweden

**Correspondence**
Tomi Myrberg, Department of Surgical and Perioperative Science, Anesthesiology and Intensive Care Medicine, Umeå University, Umeå 901 85, Sweden.
Email: tomi.myrberg@umu.se

## Abstract

**Background:** Restrictive lung function may indicate various underlying diseases. The aim of this study was to evaluate the accuracy of different restrictive spirometry patterns (RSPs) to identify restrictive lung function (total lung capacity [TLC] < lower limit of normal [LLN]) according to reference values by the Global Lung Function Initiative (GLI) in a wide age-ranged, general population sample.

**Methods:** A general population sample ($n = 607$, age 23–72 years, smokers 18.8%) with proper dynamic spirometry and TLC measurements, was included. Accuracy of two main categories of RSP to identify TLC < LLN were evaluated: traditional RSPs (definition 1: FVC < 80% of predicted and $FEV_1/FVC \geq 0.7$ and definition 2: FVC < LLN and $FEV_1/FVC \geq$ LLN) and RSPs defined by Youden's method (definition 3: FVC < 85.5% of predicted and $FEV_1/FVC \geq$ LLN and definition 4: FVC *Z*-score < −1.0 and $FEV_1/FVC \geq$ LLN).

**Results:** The prevalence of restrictive lung function (TLC < LLN) was 5.3%. The most accurate cut-offs for FVC to identify TLC < LLN were 85.5% for FVC% of predicted, and −1.0 for FVC *Z*-score. The traditional RSP definitions 1 and 2 had higher specificity (95.0% and 96.9%) but substantially lower sensitivity compared to RSP definitions 3 and 4.

**Conclusion:** Based on the GLI reference values, the RSP definition FVC < LLN and $FEV_1/FVC \geq$ LLN yielded the highest specificity and may appropriately be used to rule out restrictive lung function. The RSP definition with the most favourable trade-off between sensitivity and specificity, FVC < 85.5% of predicted and $FEV_1/FVC \geq$ LLN, may serve as an alternative with higher sensitivity for screening.

KEYWORDS

epidemiology, respiratory function tests, restrictive lung function, restrictive spirometry pattern, spirometry, total lung capacity

## 1 | INTRODUCTION

Restrictive lung function may indicate various underlying diseases that may decrease quality of life and life expectancy (Eriksson et al., 2013; Godfrey & Jankowich, 2016). Restrictive lung function is common in several conditions and diseases including lung and pleural diseases but also obesity, deformities of thorax and neuromuscular disease, and notably, other conditions than lung and pleural (Bradley et al., 2008; Stansbury & Mannino, 2009). In addition, the COVID-19 pandemic may increase the global prevalence of restrictive lung function due to inflammatory response in the lungs (E. et al., 2021; Iversen et al., 2022; Torres-Castro et al., 2021). On a population level, prevalence estimates of restrictive lung function using dynamic spirometry, that is, restrictive spirometry pattern (RSP), based on normal forced expiratory volume in a one second ($FEV_1$)/forced vital capacity (FVC) (or VC) ratios and decreased FVC (or VC) have yielded results between 6% and 8% in the United States (Ford et al., 2013; Kurth & Hnizdo, 2015), from 5% to 19% in Spain (Scarlata et al., 2008), and in Sweden about 10% (Backman et al., 2016).

Decreased total lung capacity (TLC) is the gold standard measure of restrictive lung function and referral to a pulmonary function laboratory is necessary in the diagnostic process (Bradley et al., 2008; Godfrey & Jankowich, 2016; Pellegrino et al., 2005; Wanger et al., 2005). In clinical practice, RSP has been used for a primary screening to increase diagnostic feasibility and reduce unnecessary lung volume testing, as advocated by the current ERS/ATS guidelines (Pellegrino et al., 2005), mostly to rule out restrictive pulmonary concerns. However, a study of patients referred to a tertiary care pulmonary function laboratory showed that only 41% with the FVC < lower limit of normal (LLN) and $FEV_1/FVC \geq LLN$ pattern also had TLC < LLN (Aaron et al., 1999). Despite several studies on RSP (Aaron et al., 1999; D'Aquino et al., 2010; Glady et al., 2003; Torén et al., 2020; Vandevoorde et al., 2008), there is no clear consensus on the most accurate definition of RSP to rule out pulmonary restriction (Godfrey & Jankowich, 2016). In addition, most previous studies in the field have applied older reference values for spirometry and lung volumes (Crapo et al., 1982; Quanjer et al., 1993; Torén et al., 2020; Vandevoorde et al., 2008) mainly in selected patient populations (Crapo et al., 1982; Quanjer et al., 1993; Vandevoorde et al., 2008). There is also only one recent population-based estimate of the prevalence of restrictive lung function based on TLC measurements, at 5.4%, however, this result was based on examinations of a population in a narrow age range (Torén et al., 2020).

The aim of this study was to evaluate the accuracy of different RSPs to identify restrictive lung function defined as TLC < LLN according to the 2021 reference values for TLC by the Global Lung Function Initiative (GLI; Hall et al., 2021). The secondary aim was to estimate the prevalence of restrictive lung function in a wide age-ranged population sample.

## 2 | METHODS

The study was conducted within the Obstructive Lung Disease in Northern Sweden (OLIN) research programme. In 1992, a general population sample of 5681 individuals in ages 20–72 years, living in Norrbotten County in Northern Sweden, was invited to participate in a postal questionnaire survey on respiratory diseases and symptoms. Of these, 4851 (85%) responded and in 1994–1995 a random sample of the responders ($n = 970$) was invited to clinical examinations (Lindberg et al., 2005). Structured interviews and lung function testing including pre- and post-bronchodilator (BD) spirometry, static lung volumes and diffusing capacity according to the American Thoracic Society (ATS) recommendations (American Thoracic Society, 1987, 1991) were conducted. Exclusion criteria for lung function testing included a recent myocardial infarction (within 1 month) or a limited ability to cooperate for other reasons (Miller et al., 2005). Overall, 664 participated in the study, among which both spirometry and static lung volumes that met appropriate quality requirements (American Thoracic Society, 1987, 1991; Wanger et al., 2005) were obtained in 607 individuals (age range 23–72 years). The ethical approval for this study was provided by the Regional Ethical Review Board at Umeå University (DNR 1991-236). The study adheres to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.

### 2.1 | Static lung volumes

All pulmonary volume measurements were obtained with the same methodology and device. Functional residual capacity (FRC) was evaluated from the volume of thoracic gas, measured in a body plethysmograph (Sensormedics Autobox 6200). Inspiratory and expiratory volumes were thereafter evaluated from the volume flow in a mass flow sensor (Sensormedics 2200). At least three sufficient measurements were required. Peak to peak variation of the oral pressure was not allowed to exceed 2 kPa. Open pressure-volume-loops were disregarded. The mean value of three FRC measurements was calculated and this value should not deviate from the single measurements with more than 5%. TLC was calculated as FRC + the highest value of inspiratory capacity (IC). One trained laboratory technician, with more than 20 years of experience in lung function testing, performed all measurements. Pre-bronchodilator values were used in this study, and the GLI 2021 reference values were applied (Hall et al., 2021).

### 2.2 | Spirometry

Flow-volume curves were recorded with a dry volume spirometer (Mijnhardt, Vicatest 5). At least three sufficient recordings were required, the highest values of FVC and $FEV_1$ were used for calculation of an $FEV_1/FVC$ ratio. The chosen values were not allowed to exceed the second highest with more than 5%, or 100 ml for values below

**TABLE 1** Overall characteristics and lung function measures among all participants and stratified by sex

| Parameter | All $n = 607$ | Men $n = 311$ | Women $n = 296$ | $p$ |
|---|---|---|---|---|
| Restrictive lung function (TLC < LLN), *n* (%) | 32 (5.3) | 15 (4.8) | 17 (5.7) | 0.612 |
| Age (years) | 48.6 ± 12.5 | 49.4 ± 12.2 | 47.8 ± 13 | 0.123 |
| Height (cm) | 170.7 ± 9.1 | 177.1 ± 6.5 | 163.8 ± 5.8 | <0.001 |
| Body mass index (kg/m$^2$) | 25.6 ± 3.5 | 26.0 ± 3.1 | 25.3 ± 4.0 | 0.013 |
| Never-smoker, *n* (%) | 269 (44) | 126 (41) | 143 (48) | 0.059 |
| Ex-smoker,[a] *n* (%) | 179 (29) | 115 (37) | 64 (22) | <0.001 |
| Smoker, *n* (%) | 159 (26) | 70 (23) | 89 (30) | 0.032 |
| Any wheeze last 12 months, *n* (%) | 213 (35) | 105 (34) | 108 (36) | 0.463 |
| TLC% of predicted | 99.2 ± 11.7 | 98.6 ± 11.1 | 99.9 ± 12.2 | 0.182 |
| TLC *Z*-score | −0.084 ± 0.99 | −0.12 ± 0.96 | −0.04 ± 1.01 | 0.307 |
| $FEV_1$% of predicted | 92.7 ± 15.6 | 92.8 ± 15.7 | 92.5 ± 15.6 | 0.795 |
| $FEV_1$ *Z*-score | −0.51 ± 1.1 | −0.49 ± 1.1 | −0.53 ± 1.1 | 0.688 |
| FVC% of predicted | 94.7 ± 13.4 | 94.1 ± 13.0 | 95.5 ± 13.7 | 0.192 |
| FVC *Z*-score | −0.38 ± 0.97 | −0.43 ± 0.95 | −0.32 ± 1.0 | 0.150 |
| $FEV_1$/FVC | 0.80 ± 0.024 | 0.79 ± 0.020 | 0.81 ± 0.023 | <0.001 |
| $FEV_1$/FVC *Z*-score | −0.29 ± 1.0 | −0.17 ± 1.0 | −0.41 ± 1.0 | 0.004 |
| $FEV_1$/FVC < LLN, *n* (%) | 56 (9) | 27 (9) | 29 (10) | 0.635 |

*Note*: Results presented as mean ± *SD* unless otherwise stated.

Abbreviations: FEV1, forced expiratory volume in 1 s; FVC, forced vital capacity; LLN, lower limit of normal; n, number; SD, standard deviation; TLC, total lung capacity.

[a]Ex-smoker = smoked for at least 1 year but not during the last 12 months.

1 L. Pre-bronchodilator values were used in this study. All spirometry measurements were performed by nurses specially trained for this purpose. The GLI 2012 reference values were applied (Quanjer et al., 2012).

## 2.3 | Definitions

The definition of decreased total lung volume, that is, a restrictive lung function, was TLC < LLN, in accordance with current guidelines (Pellegrino et al., 2005), with the LLN defined as the fifth percentile of the GLI reference values (Hall et al., 2021). Accuracy of two main categories of RSPs to identify TLC < LLN were evaluated: (a) traditional, that is, commonly used in epidemiological studies (Godfrey & Jankowich, 2016), and European Respiratory Society/American Thoracic Society (ERS/ATS) guideline-suggested (Pellegrino et al., 2005) RSPs and (b) RSPs defined by Youden's method (Youden, 1950) based on the study sample. The traditional RSPs (Godfrey & Jankowich, 2016) were defined as a combination of decreased FVC and a normal or increased $FEV_1$/FVC ratio as follows:

**Definition 1.** (Godfrey & Jankowich, 2016): FVC < 80% of predicted and $FEV_1$/FVC ≥ 0.7.

**Definition 2.** (Pellegrino et al., 2005): FVC < LLN and $FEV_1$/FVC ≥ LLN.

The RSPs defined by Youden's method were defined based on cut-offs with the highest Youden index for the spirometry measures in the study sample as follows:

**Definition 3.** FVC < [new cut-off for FVC% of predicted] and $FEV_1$/FVC ≥ LLN.

**Definition 4.** FVC < [new cut-off for FVC Z-score] and $FEV_1$/FVC ≥ LLN.

## 2.4 | Statistics

The study data were analysed with SPSS (IBM SPSS Statistics for Macintosh, Version 26.0. Armonk, NY: IBM Corp). Chi-square test was used to compare proportions, and Student's *t*-test to compare means. $p$ values < 0.05 from two-sided tests were considered statistically significant. Sensitivity, specificity and positive and negative predictive values were calculated for the respective RSPs ability to identify TLC < LLN. The equation used to calculate test efficiency

was: (∑true positive cases + ∑true negative cases)/∑overall cases. Receiver operating characteristic (ROC) curves and area under the curves (AUROC) were calculated to determine accuracy. The most accurate cut-off values for the two new RSPs were defined with Youden's method (Youden, 1950) from the coordinate tables for respective AUROC. Youden's method (Youden index = sensitivity + specificity − 1) is an acknowledged way to define the highest combination of sensitivity and specificity for a cut-off for a specific variable.

Sensitivity analysis was conducted by use of VC, that is, the highest of forced and slow vital capacity, instead of FVC.

## 3 | RESULTS

The prevalence of restrictive lung function defined as TLC < LLN was 5.3%, and there were no differences between men and women (Table 1). Basic characteristics did not differ significantly between the groups with and without restrictive lung function (Table 2). Spirometry values were lower in individuals with restrictive lung function versus those without, with $FEV_1$ percent of predicted 77.4% versus 93.5%, $p < 0.001$, and FVC% of predicted 76.3% versus 95.8%, $p < 0.001$. Mean TLC% of predicted and *Z*-score was 100.5% and 0.03, respectively, among those without restrictive lung function (Table 2). In addition, mean (±*SD*) TLC *Z*-scores and TLC% of predicted were −0.18 ± 0.92 and 97.6 ± 12.8, respectively, in never-smokers with no self-reported or doctor-diagnosed asthma, chronic bronchitis or emphysema (Figure 1). There were no differences in TLC or spirometry values expressed as per cent of predicted, or *Z*-scores, by sex.

The most accurate cut-offs for RSPs defined by Youden's method were for FVC% of predicted 85.5%, FVC *Z*-score −1.0, $FEV_1$% of predicted 87.5%, and $FEV_1$ *Z*-score −0.98. Of these, FVC *Z*-score (cut-off −1.0) and FVC% of predicted (cut-off 85.5%) had superior sensitivity, specificity and AUROC with very narrow 95% confidence intervals compared to both per cent of predicted and *Z*-score for $FEV_1$ (Figures 2 and 3 and Table 3). In addition, $FEV_1$/FVC ratio (cut-off 79.3%) and $FEV_1$/FVC *Z*-score (cut-off −0.457) had both low AUROC (0.597 and 0.615, $p = 0.064$ and $p = 0.029$, respectively). The cut-offs with best accuracy for identifying TLC < LLN (i.e., FVC% of predicted and FVC *Z*-score) were applied in the RSP Definitions 3 and 4 (Table 4).

The prevalence of RSP Definitions 1–4 was 7.4%, 5.4%, 18.5% and 20.6%, respectively. The traditional RSP Definitions 1 and 2 had higher specificity, but substantially lower sensitivity compared to the RSP Definitions 3 and 4 by Youden's method (Table 4). Among the single spirometry measures, FVC < 85.5% and FVC *Z*-score < −1.0 yielded equivalent negative predictive values but slightly lower test efficiency compared to RSP Definitions 1–4 (Tables 3 and 4). However, the traditional RSP Definitions 1 and 2 had higher specificity (i.e., ability to rule out the disease) compared to all other RSPs (Table 4). The overlap of the different RSP definitions and restrictive lung function is illustrated by proportional Venn diagrams (Figures 4 and S1). Similar results regarding accuracy and prevalence of RSP Definitions 1–4 were found by using VC instead of FVC. This data is presented in Supporting Information file 2.

**TABLE 2** Characteristics and lung function measures in individuals with and without restrictive lung function defined by TLC < LLN

| Parameter | Restrictive lung function, *n* = 32 | No restrictive lung function, *n* = 575 | *p* |
|---|---|---|---|
| Age | 51.4 ± 11.6 | 48.4 ± 12.6 | 0.187 |
| Female sex, *n* (%) | 17 (53.1) | 279 (48.5) | 0.612 |
| Height | 168.5 ± 8.8 | 170.8 ± 9.1 | 0.166 |
| Body mass index (kg/m$^2$) | 26.1 ± 4.2 | 25.6 ± 3.5 | 0.406 |
| Never-smoker, *n* (%) | 15 (46.9) | 254 (44.2) | 0.756 |
| Ex-smoker,[a] *n* (%) | 11 (34.4) | 168 (29.2) | 0.538 |
| Smoker, *n* (%) | 6 (18.8) | 153 (26.6) | 0.323 |
| PD asthma | 6 (18.8) | 55 (9.6) | 0.093 |
| PD emphysema or chronic bronchitis | 2 (6.3%) | 31 (5.4%) | 0.326 |
| TLC% of predicted | 76.7 ± 5.2 | 100.5 ± 10.6 | <0.001 |
| TLC *Z*-score | −2.1 ± 0.5 | 0.03 ± 0.9 | <0.001 |
| $FEV_1$% of predicted | 77.4 ± 13.3 | 93.5 ± 15.3 | <0.001 |
| $FEV_1$ *Z*-score | −1.6 ± 0.9 | −0.4 ± 1.1 | <0.001 |
| FVC% of predicted | 76.3 ± 10.7 | 95.8 ± 12.7 | <0.001 |
| FVC *Z*-score | −1.7 ± 0.8 | −0.3 ± 0.9 | <0.001 |
| $FEV_1$/FVC | 0.8 ± 0.02 | 0.8 ± 0.02 | 0.338 |
| $FEV_1$/FVC *Z*-score | 0.1 ± 0.8 | −0.3 ± 1.0 | 0.020 |

*Note*: Results presented as number and proportion (%) of individuals and/or mean ± *SD* unless otherwise stated.

Abbreviations: FEV1, forced expiratory volume in 1 s; FVC, forced vital capacity; LLN, lower limit of normal; PD, physician diagnosis; SD, standard deviation; TLC, total lung capacity.

[a]Ex-smoker = smoked for at least 1 year but not during the last 12 months.

## 4 | DISCUSSION

The accuracy of four definitions of RSP to identify restrictive lung function (TLC < LLN) according to the 2021 GLI reference values were evaluated in this population-based study (Hall et al., 2021). The ERS/ATS guideline supported definition for RSP of FVC < LLN and $FEV_1$/FVC ≥ LLN had the best ability to rule out those not having restrictive lung function, while FVC < 85.5% and $FEV_1$/FVC ≥ LLN had the best ability to identify those with restrictive lung function. Moreover, the prevalence of restrictive lung

function was 5.3% in our adult sample. Importantly, TLC was clearly centred around 100% of predicted and *Z*-score 0 in those without restrictive lung function, indicating an appropriate fit for the 2021 GLI reference values for static lung volumes in this Swedish population.

The ERS/ATS guideline suggests using the LLN for defining a normal $FEV_1/VC$ ratio along with low VC as primary screening for restrictive lung function (Pellegrino et al., 2005). As FVC today is the most commonly assessed measure of vital capacity, and as it performs similarly as VC in this setting (Glady et al., 2003), we used FVC instead of VC in our definitions of RSP, in line with most other studies (Aaron et al., 1999; Ford et al., 2013; Glady et al., 2003; Godfrey & Jankowich, 2016; Kurth & Hnizdo, 2015; Torén et al., 2020). Of importance, the ERS/ATS guideline supported definition (Pellegrino et al., 2005) of RSP yielded a similar prevalence as observed regarding TLC < LLN, that is, 5.4% compared to 5.3%, as well as compared to a 5.4% prevalence of TLC < LLN in another Swedish study (Torén et al., 2020). In contrast, the two RSPs defined by Youden's method both overestimated the prevalence of restrictive lung function almost fourfold, a prevalence substantially higher also compared with previous reports on RSP prevalence (Backman et al., 2016; Torén et al., 2020). In line with results from others (Aaron et al., 1999), all the combined RSP definitions, that is, including also the $FEV_1/FVC$ ratio, are preferable to single spirometry values to detect signs of restrictive lung function, as both FVC% of predicted and FVC *Z*-score as single measures had slightly lower test efficiency and substantially lower specificity compared to the combined RSPs.

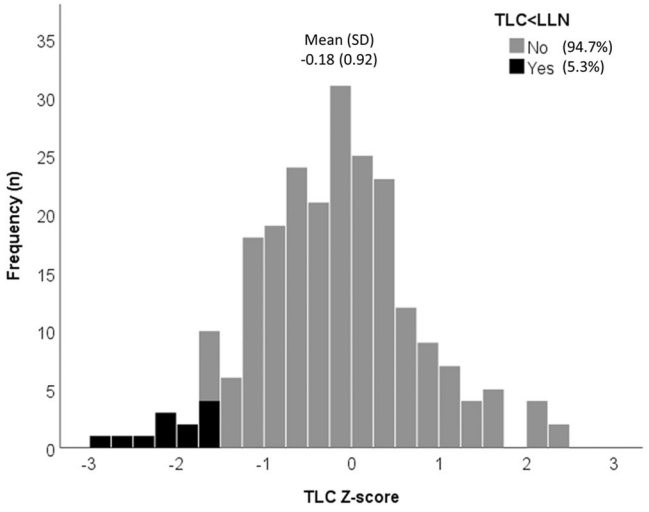


**FIGURE 1** Histogram of frequency distribution for total lung capacity (TLC) *Z*-scores in never-smokers with no self-reported or doctor-diagnosed asthma, chronic bronchitis or emphysema ($n = 228$). Grey bars indicate TLC $Z \geq$ LLN while black bars indicate TLC $Z <$ LLN. LLN, lower limit of normal

The lower sensitivity of the traditional definitions of RSPs is explained by lower cut-off for FVC < 80% of predicted compared to FVC < 85.5% of predicted defined by Youden´s method. This is well in line with previous findings based on patient-based materials (Glady et al., 2003) that cut-offs around 85% for FVC % of predicted may be both most accurate and should be applied in a clinical setting instead of the conventional cut-off at 80% of predicted. In line with our results, other studies have indicated that about half or less of patients with FVC < LLN and $FEV_1/FVC \geq$ LLN (Aaron et al., 1999) in fact have TLC < LLN, and similar results have been seen for the RSP FVC < 85% and $FEV_1/FVC \geq 0.55$ (Glady et al., 2003). Importantly, according to the 2021 GLI reference values (Hall et al., 2021), FVC% of predicted <85.5% or FVC *Z*-score < −1.0 performed similarly well regarding ability to identify restrictive lung function. Additionally, FVC performed similarly as VC, that is, no noteworthy differences between FVC and VC derived RSPs regarding prevalence or accuracy to identify TLC < LLN were found in this cohort.

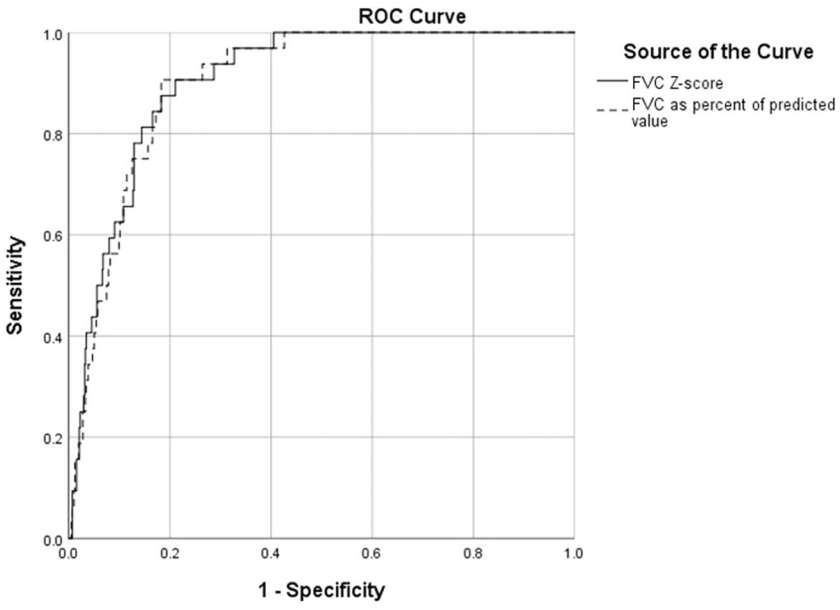


**FIGURE 2** Receiver operating characteristic (ROC) curves for the ability of FVC *Z*-score (AUROC: 0.904, 95% CI: 0.867–0.941) and FVC% of predicted (AUROC: 0.900, 95% CI: 0.863-0.937) to identify restrictive lung function (TLC < lower limit of normal)

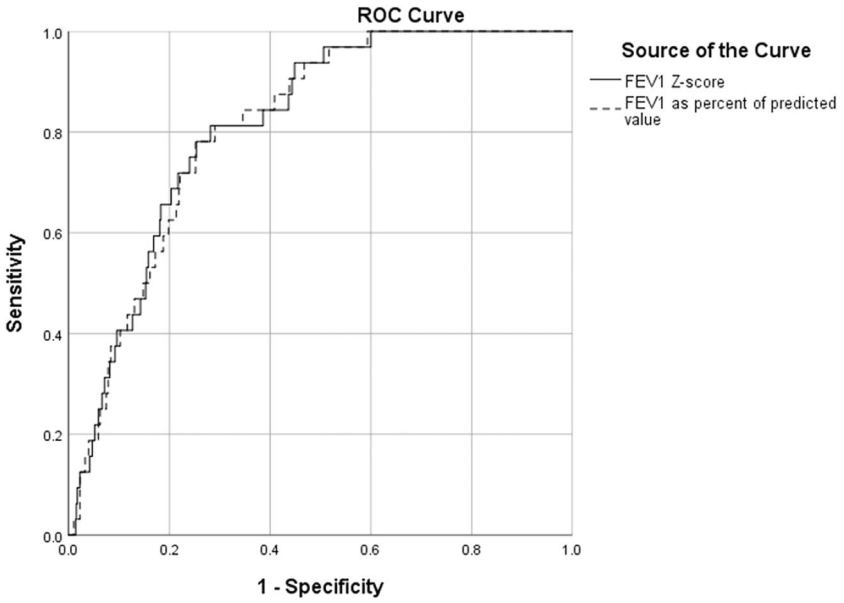


**FIGURE 3** Receiver operating characteristic (ROC) curves for the ability of $FEV_1$ *Z*-score (AUROC: 0.814, 95% CI: 0.755–0.872) and $FEV_1$% of predicted (AUROC: 0.812, 95% CI: 0.754–0.870) to identify restrictive lung function (TLC < lower limit of normal)

**TABLE 3** Accuracy for spirometry measures to discriminate restrictive lung function (TLC < LLN)

| Parameter | Sensitivity (%) | Specificity (%) | AUROC (95% CI) | *p* |
|---|---|---|---|---|
| FVC *Z*-score (cut-off −1.0) | 87.5 | 79.1 | 0.904 (0.867–0.941) | <0.001 |
| FVC% of predicted (cut-off 85.5%) | 90.6 | 81.7 | 0.900 (0.863–0.937) | <0.001 |
| $FEV_1$ *Z*-score (cut-off −0.98) | 81.3 | 71.8 | 0.814 (0.755–0.872) | <0.001 |
| $FEV_1$% of predicted (cut-off 87.5%) | 81.3 | 71.0 | 0.812 (0.754–0.870) | <0.001 |
| $FEV_1$/FVC (cut-off 79.3%) | 65.6 | 52.9 | 0.597 (0.505–0.689) | 0.064 |
| $FEV_1$/FVC *Z*-score (cut-off −0.457) | 71.9 | 39.8 | 0.615 (0.519–0.710) | 0.029 |

*Note*: The optimal cut-off values, sensitivity and specificity for respective parameter are defined by Youden's method (i.e., highest Youden's index).
Abbreviations: AUROC, area under the ROC curve; CI, confidence interval; FEV1, forced expiratory volume in 1 s; FVC, forced vital capacity; LLN, lower limit of normal; RSP, restrictive spirometry pattern; TLC, total lung capacity; Z-score, standardized residual.

**TABLE 4** Overall performance and accuracy of different restrictive spirometry patterns (RSP) to discriminate restrictive lung function (TLC < LLN)

| RSP | Efficiency | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) |
|---|---|---|---|---|---|
| ***Definition 1***: FVC < 80% of pred and $FEV_1$/FVC ≥ 0.7 | 0.93 | 50.0 | 95.0 | 35.6 | 97.2 |
| ***Definition 2***: FVC < LLN and $FEV_1$/FVC ≥ LLN | 0.94 | 46.9 | 96.9 | 45.5 | 97.0 |
| ***Definition 3***: FVC < 85.5% of pred[a] and $FEV_1$/FVC ≥ LLN | 0.86 | 90.6 | 85.6 | 25.9 | 99.4 |
| ***Definition 4***: FVC Z-score < −1.0[a] and $FEV_1$/FVC ≥ LLN | 0.83 | 87.5 | 83.1 | 24.4 | 99.2 |
| FVC < 85.5% of pred[a] | 0.82 | 90.6 | 81.7 | 21.6 | 99.4 |
| FVC Z-score < −1.0[a] | 0.80 | 87.5 | 79.1 | 18.9 | 99.1 |

*Note*: The prevalence of RSP Definitions 1–4 was 7.4% (45/607), 5.4% (33/607), 18.5% (112/607) and 20.6% (125/607), respectively.
Abbreviations: FEV1, forced expiratory volume in 1 s; FVC, forced vital capacity; LLN, lower limit of normal; NPV, negative predictive value; PPV, positive predictive value; RSP, restrictive spirometry pattern; TLC, total lung capacity, % of pred, percent of predicted value.
[a]The best cut-off defined by Youden's method. PPV, NPV, sensitivity and specificity were calculated from cross-tabulations.

It is essential to take into account the importance of sensitivity and specificity in relation to the purpose of performing the test in terms of 'ruling out' or 'overall screening'. When assessing an individual patient, it is important to prioritize specificity if the aim is to accurately rule out those not having restrictive lung function and avoid unnecessary referral to pulmonary function laboratory for testing of static lung volumes. In many epidemiologic and clinical screening studies, identification of only 'true' cases with a disease

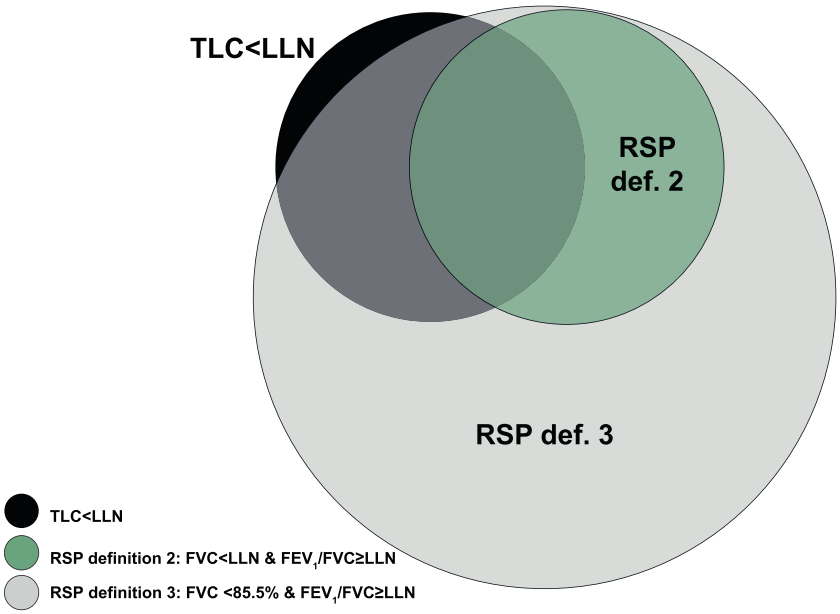


**FIGURE 4** Proportional Venn diagrams for overlap between two different restrictive spirometry patterns (RSPs) and restrictive lung function (TLC < lower limit of normal)

may be most important, that is, to accurately rule out those without the disease. In comparison, when utilizing overall screening and assessing population prevalence, the trade-off between sensitivity and specificity may be most important. Sensitivity should be prioritized in settings where it is imperative to find all cases, for example, in a setting with highly contagious diseases with severe outcome. Somewhat contrasting, Glady et al. (2003) emphasized the importance of sensitivity rather than specificity of RSP in screening for restrictive lung function. Especially in rural or geographically large areas, as our study area, travelling logistics to and from the pulmonary function laboratory may add additional costs on top of those for the testing. Nevertheless, regardless of screening test and approach used, for a definitive diagnosis of restrictive lung function, a measurement of TLC is needed (Aaron et al., 1999; Crapo, 1994; Pellegrino, 2005).

For a proper diagnostic accuracy, there is a need for reference values estimated with state-of-the-art methods that cover wide age spans. An important step is to make sure that the selected reference values fit the population under study as this undoubtedly will affect the outcome (Backman et al., 2015). In Sweden, the Hedenström reference values (Hedenstrom et al., 1985, 1986), a Swedish standard material published in the 1980s including dynamic spirometry, static lung volumes and diffusing capacity, has been commonly used. ERS Task Forces have contributed with the GLI multiethnic, wide age-ranged reference values for spirometry published 2012 (Quanjer et al., 2012) and now for static lung volumes in 2021 (Hall et al., 2021) for individuals of European ancestry. The GLI 2021 reference material for lung volumes does not include Swedish data, but importantly, our results indicated an appropriate fit for this Swedish sample. Notably, the % predicted values did not quite align in this sample when using the GLI reference values (Quanjer et al., 2012) for FEV1 and FVC (93% and 95% of predicted) and the GLI reference values (Hall et al., 2021) for TLC (99% of predicted). Previous results have also revealed some weaknesses regarding fit for the GLI reference values for spirometry (Quanjer et al., 2012) in the Swedish population (Backman et al., 2015). This should be taken into consideration when interpreting spirometry results in the context of RSPs.

## 5 | LIMITATIONS AND STRENGTHS OF THE STUDY

Due to relatively low prevalence, the absolute number of cases with TLC < LLN was low, and thus some bias in the estimates of sensitivity and specificity for the RSPs cannot be ruled out. However, as the results indicated proper fit for the 2021 GLI reference values for static lung volumes, the prevalence estimates should be considered reliable and generalizable. Further, the participation rate was high, and a nonresponder study performed within OLIN indicated only limited or no bias (Raisanen et al., 2020). To our knowledge, no previous studies in this study field have utilized established methods by Youden to identify the most favourable cut-off for FVC and VC. In addition, repeated method quality controls were performed throughout the study and the spirometers were calibrated every morning on each study day. Finally, despite being performed according to the guidelines in practice at the time of data collection, we acknowledge that a more recent study potentially could find slightly different prevalence estimates, for example, due to different population characteristics in terms of smoking habits, diet, sedentary lifestyle and BMI.

## 6 | CONCLUSION

Based on the GLI reference values for lung volumes, the prevalence of restrictive lung function defined as TLC < LLN was 5.3%. The ERS/ATS guideline-defined RSP (FVC < LLN and $FEV_1/FVC \geq$ LLN) yielded

the highest test efficiency and specificity. The RSP with the most favourable trade-off between sensitivity and specificity (FVC < 85% of predicted and $FEV_1/FVC \geq LLN$) may serve alternative possibilities with higher sensitivity for screening of restrictive lung function. Furthermore, the results indicated a proper fit for the GLI reference values for static lung volumes in this Swedish population.

## ACKNOWLEDGEMENTS

This study is dedicated to the memory of professor Staffan Andersson, rest in peace. All research staff is acknowledged for data collection, especially Ann-Christin Skogsberg, Ulla Jarlbring, and Ann-Christine Jonsson.

## CONFLICT OF INTERESTS

BL has received grants from Astra Zeneca for scientific work outside this submitted manuscript, and a personal fee from Sanofi for participation to an advisory board. CS has received payments from Astra Zeneca, Boehringer Ingelheim, and Novartis for educational lectures and participated on an advisory board of Novartis. AL has received personal fees from Boehringer Ingelheim and Novartis for lectures at educational events, and personal fees from Astra Zeneca, GlaxoSmithKline, Novartis, and Boehringer Ingelheim for participation on advisory boards. HB has received payments from Astra Zeneca and Boehringer Ingelheim for presentation at scientific meetings outside this submitted manuscript. All other authors declare no conflict of interests.

## AUTHOR CONTRIBUTIONS

Tomi Myrberg: Conceptualization, Methodology, Formal analysis, Writing–original draft, Writing–review & editing, Visualization. Anne Lindberg: Writing–review & editing, Visualization, Supervision. Berne Eriksson: Writing–review & editing, Visualization, Formal analysis. Linnea Hedman: Writing–review & editing, Visualization. Caroline Stridsman: Writing–review & editing, Visualization. Bo Lundbäck: Writing–review & editing, Supervision. Eva Rönmark: Writing–review & editing, Supervision. Helena Backman: Conceptualization, Methodology, Formal analysis, Writing–review & editing, Visualization, Data curation, Project administration.

## DATA AVAILABILITY STATEMENT

Data are available on a reasonable request.

## ORCID

*Tomi Myrberg* https://orcid.org/0000-0002-8802-2321

## REFERENCES

Aaron, S.D., Dales, R.E. & Cardinal, P. (1999) How accurate is spirometry at predicting restrictive pulmonary impairment? *Chest*, 115(3), 869–873.

American Thoracic Society. (1987) Standardization of spirometry—1987 update. Statement of the American Thoracic Society. *The American Review of Respiratory Disease*, 136(5), 1285–1298.

American Thoracic Society. (1991) Lung function testing: selection of reference values and interpretative strategies. American Thoracic Society. *The American Review of Respiratory Disease*, 144(5), 1202–1218.

Backman, H., Lindberg, A., Sovijarvi, A., Larsson, K., Lundback, B. & Ronmark, E. (2015) Evaluation of the global lung function initiative 2012 reference values for spirometry in a Swedish population sample. *BMC Pulmonary Medicine*, 15, 26.

Backman, H., Eriksson, B., Hedman, L., Stridsman, C., Jansson, S.A. & Sovijärvi, A. et al. (2016) Restrictive spirometric pattern in the general adult population: methods of defining the condition and consequences on prevalence. *Respiratory Medicine*, 120, 116–123.

Bradley, B., Branley, H.M., Egan, J.J., Greaves, M.S., Hansell, D.M. & Harrison, N.K. et al. (2008) Interstitial lung disease guideline: the British Thoracic Society in collaboration with the Thoracic Society of Australia and New Zealand and the Irish Thoracic Society. *Thorax*, 63(Suppl 5), v1–v58.

Crapo, R.O. (1994) Pulmonary-function testing. *The New England Journal of Medicine*, 331(1), 25–30.

Crapo, R.O., Morris, A.H., Clayton, P.D. & Nixon, C.R. (1982) Lung volumes in healthy nonsmoking adults. *Bulletin Europeen de Physiopathologie Respiratoire*, 18(3), 419–425.

D'Aquino, L.C., Rodrigues, S.C., Barros, J.A., Rubin, A.S., Rosário Filho, N.A. & Pereira, C.A. (2010) Predicting reduced TLC in patients with low FVC and a normal or elevated FEV1/FVC ratio. *Jornal Brasileiro de Pneumologia*, 36(4), 460–467.

E, E., R, F., Öi, E., Im, L., M, L. & S, R. et al. (2021) Impaired diffusing capacity for carbon monoxide is common in critically ill Covid-19 patients at four months post-discharge. *Respiratory Medicine*, 182, 106394.

Eriksson, B., Lindberg, A., Mullerova, H., Ronmark, E. & Lundback, B. (2013) Association of heart diseases with COPD and restrictive lung function—results from a population survey. *Respiratory Medicine*, 107(1), 98–106.

Ford, E.S., Mannino, D.M., Wheaton, A.G., Giles, W.H., Presley-Cantrell, L. & Croft, J.B. (2013) Trends in the prevalence of obstructive and restrictive lung function among adults in the United States: findings from the National Health and Nutrition Examination surveys from 1988-1994 to 2007-2010. *Chest*, 143(5), 1395–1406.

Glady, C.A., Aaron, S.D., Lunau, M., Clinch, J. & Dales, R.E. (2003) A spirometry-based algorithm to direct lung function testing in the pulmonary function laboratory. *Chest*, 123(6), 1939–1946.

Godfrey, M.S. & Jankowich, M.D. (2016) The vital capacity is vital: epidemiology and clinical significance of the restrictive spirometry pattern. *Chest*, 149(1), 238–251.

Hall, G.L., Filipow, N., Ruppel, G., Okitika, T., Thompson, B. & Kirkby, J. et al. (2021) Official ERS technical standard: global lung function initiative reference values for static lung volumes in individuals of European ancestry. *The European Respiratory Journal*, 57(3), 2000289.

Hedenstrom, H., Malmberg, P. & Agarwal, K. (1985) Reference values for lung function tests in females. Regression equations with smoking variables. *Bulletin Europeen de Physiopathologie Respiratoire*, 21(6), 551–557.

Hedenstrom, H., Malmberg, P. & Fridriksson, H.V. (1986) Reference values for lung function tests in men: regression equations with smoking variables. *Upsala Journal of Medical Sciences*, 91(3), 299–310.

Iversen, K.K., Afzal, S., Ahlström, M.G., Nordestgaard, B.G., Schneider, U.V. & Nielsen, L. et al. (2022) Lung function decline in relation to COVID-19 in the general population: a matched cohort study with pre-pandemic assessment of lung function. *The Journal of Infectious Diseases*, jiab636.

Kurth, L. & Hnizdo, E. (2015) Change in prevalence of restrictive lung impairment in the U.S. population and associated risk factors: the National Health and Nutrition Examination Survey (NHANES) 1988-1994 and 2007-2010. *Multidisciplinary Respiratory Medicine*, 10(1), 7.

Lindberg, A., Jonsson, A.C., Ronmark, E., Lundgren, R., Larsson, L.G. & Lundback, B. (2005) Prevalence of chronic obstructive pulmonary disease according to BTS, ERS, GOLD and ATS criteria in relation to doctor's diagnosis, symptoms, age, gender, and smoking habits. *Respiration*, 72(5), 471–479.

Miller, M.R., Crapo, R., Hankinson, J., Brusasco, V., Burgos, F. & Casaburi, R. et al. (2005) General considerations for lung function testing. *The European Respiratory Journal*, 26(1), 153–161.

Pellegrino, R. (2005) Interpretative strategies for lung function tests. *The European Respiratory Journal*, 26(5), 948–968.

Quanjer, P.H., Tammeling, G.J., Cotes, J.E., Pedersen, O.F., Peslin, R. & Yernault, J.C. (1993) Lung volumes and forced ventilatory flows. Report working party standardization of lung function tests, European community for steel and coal. Official statement of the European Respiratory Society. *European Respiratory Journal*, 16, 5–40.

Quanjer, P.H., Stanojevic, S., Cole, T.J., Baur, X., Hall, G.L. & Culver, B.H. et al. (2012) Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *European Respiratory Journal*, 40(6), 1324–1343.

Raisanen, P., Hedman, L., Andersson, M., Stridsman, C., Lindberg, A. & Lundback, B. et al. (2020) Non-response did not affect prevalence estimates of asthma and respiratory symptoms–results from a postal questionnaire survey of the general population. *Respiratory Medicine*, 173, 106017.

Scarlata, S., Pedone, C., Fimognari, F.L., Bellia, V., Forastiere, F. & Incalzi, R.A. (2008) Restrictive pulmonary dysfunction at spirometry and mortality in the elderly. *Respiratory Medicine*, 102(9), 1349–1354.

Stansbury, R. & Mannino, D. (2009) *Diseases associated with restrictive lung function impairment*. European Respiratory Society Monographs, pp. 142–149.

Torén, K., Schiöler, L., Brisman, J., Malinovschi, A., Olin, A.C. & Bergström, G. et al. (2020) Restrictive spirometric pattern and true pulmonary restriction in a general population sample aged 50-64 years. *BMC Pulmonary Medicine*, 20(1), 55.

Torres-Castro, R., Vasconcello-Castillo, L., Alsina-Restoy, X., Solis-Navarro, L., Burgos, F., & Puppo, H. et al. (2021) Respiratory function in patients post-infection by COVID-19: a systematic review and meta-analysis. *Pulmonology*, 27(4), 328–337.

Vandevoorde, J., Verbanck, S., Schuermans, D., Broekaert, L., Devroey, D. & Kartounian, J. et al. (2008) Forced vital capacity and forced expiratory volume in six seconds as predictors of reduced total lung capacity. *European Respiratory Journal*, 31(2), 391–395.

Wanger, J., Clausen, J.L., Coates, A., Pedersen, O.F., Brusasco, V. & Burgos, F. et al. (2005) Standardisation of the measurement of lung volumes. *European Respiratory Journal*, 26(3), 511–522.

Youden, W.J. (1950) Index for rating diagnostic tests. *Cancer*, 3(1), 32–35.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Myrberg, T., Lindberg, A., Eriksson, B., Hedman, L., Stridsman, C., Lundbäck, B. et al. (2022) Restrictive spirometry versus restrictive lung function using the GLI reference values. *Clinical Physiology and Functional Imaging*, 42, 181–189. https://doi.org/10.1111/cpf.12745