



Predicted values for the forced expiratory flow adjusted for forced vital capacity, a descriptive study

Claire A. Cox ^{1,2}, Judith M. Vonk ^{2,3}, Huib A.M. Kerstjens ^{1,2},
Maarten van den Berge^{1,2} and Nick H.T. ten Hacken^{1,2}

Affiliations: ¹University of Groningen, University Medical Centre Groningen, Dept of Pulmonary Diseases Groningen, The Netherlands. ²University of Groningen, University Medical Centre Groningen, Groningen Research Institute for Asthma and COPD, Groningen, The Netherlands. ³University of Groningen, Dept of Epidemiology, Groningen, The Netherlands.

Correspondence: Claire A. Cox, Universitair Medisch Centrum Groningen, Afdeling Longziekten, HPC: AA11, Postbus 30.001, 9700 RB Groningen, the Netherlands. E-mail: c.a.cox@umcg.nl

ABSTRACT

Background: The forced expiratory flows (FEFs) towards the end of the expiration may be more sensitive in detecting peripheral airways obstruction compared to the forced expiratory volume in 1 s and forced vital capacity (FVC). However, they are highly variable. A partial solution is to adjust the FEFs for FVC (FEF/FVC). Here we provide reference equations for these adjusted FEFs at 25%, 50%, 75% and 25–75% of FVC, which are currently lacking.

Methods: We included pulmonary healthy, never-smoker adults; 14 472 subjects from Lifelines, a biobank for health research, and 338 subjects from the department's control cohorts (NORM and Fiddle). Reference equations were obtained by linear regression on 80% of the Lifelines dataset and validated on the remaining data. The best model was defined as the one with the highest adjusted R²-value. The difference in variability between adjusted and unadjusted FEFs was evaluated using the coefficient of variation.

Results: For all adjusted FEFs, the best model contained age, height and weight. The adjustment improved the coefficient of variation of the FEF₇₅ from 39% to 36% and from 43% to 40%, respectively, in males and females. The highest percentage of explained variance by the reference equation was obtained for FEF₇₅/FVC, 32%–38% for males, and 41%–46% for females, depending on the validation set.

Conclusion: We developed reference equations for FVC-adjusted FEF values. We demonstrated minimally yet significantly improved variability. Future studies in obstructive airway diseases should demonstrate whether it is worthwhile to use these (predicted) adjusted FEF values.



@ERSpublications

FEFs may be more sensitive in detecting peripheral airway obstruction compared to FEV₁ or FVC. However, they are highly variable. Adjusting the FEF by dividing it by FVC may partially solve this. This article provides FVC-adjusted FEF reference equations. <https://bit.ly/2IG2fsx>

Cite this article as: Cox CA, Vonk JM, Kerstjens HAM, *et al.* Predicted values for the forced expiratory flow adjusted for forced vital capacity, a descriptive study. *ERJ Open Res* 2020; 6: 00426-2020 [<https://doi.org/10.1183/23120541.00426-2020>].

This article has supplementary material available from openres.ersjournals.com

For this research, the data of three cohorts is used. The Lifelines biobank was registered in the Dutch Central Committee of Human Research register (CCMO-register) under number NL17981.042.07. The NORM study was registered at clinicaltrials.gov with number NCT00848406. The Fiddle study was registered at the research office of the University Medical Center Groningen (UMCG) under number 201501210.

Data availability: The data of the NORM and Fiddle cohort analysed during the current study are available from the corresponding author on reasonable request. Data from the Lifelines cohort are partly restricted (see www.lifelines.nl/researcher/).

Received: 24 June 2020 | Accepted after revision: 9 Oct 2020

Copyright ©ERS 2020. This article is open access and distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0.



Introduction

Worldwide, spirometry is the most frequently used pulmonary function test, with the main goal to assess expiratory airflow obstruction in chronic diseases, such as asthma and COPD. Airflow obstruction is usually assessed with the ratio between the forced expiratory volume in 1 s (FEV_1) and the forced vital capacity (FVC) in combination with flow–volume curves. These flow–volume curves are visually attractive and offer pattern recognition in certain situations [1]. They also allow visual representation of the adequacy of a subject's effort in the early, mid and late phase of forced expiration. Unfortunately, the corresponding numerical values for the forced expiratory flows at 25%, 50%, 75% and 25–75% of the FVC (FEF_{25} , FEF_{50} , FEF_{75} , and FEF_{25-75} , respectively) demonstrate considerable variability in the healthy population [1, 2]. Therefore, the global standards use the more reproducible and well-defined reference values of FEV_1 and (F)VC to define, grade and monitor airflow obstruction [3, 4]. However, the FEV_1 is deemed not a sensitive parameter to detect small airways disease, as the volume and flow rate of exhaled air in the first second of expiration depends mainly on the diameter and resistance of the large airways. In contrast, the FEFs, towards the end of the expiration, are more sensitive to peripheral airway narrowing, so it would be worthwhile to reduce their variability.

One important source of the high variability in the FEFs originates from their dependency on the FVC. By definition, the FEF_{25} , FEF_{50} , FEF_{75} and FEF_{25-75} values depend on FVC, so small changes in FVC may translate in considerable changes in FEF values. In clinical practice, this may have important consequences. For example, if a patient shows a good response in both flow and volume on a bronchodilator, a positive effect on flow at a certain percentage of the FVC may be underestimated due to a higher FVC. The variability of the FEF values can be reduced by adjusting the FEF values for FVC (FEF/FVC). In 1974, GREEN *et al.* [5] described this calculation when they tried to reduce the large intersubject variability of flow, by what they called size compensation. In their opinion, the uneven growth between lung size and airway calibre (called dysanapsis) was an important contributor to intersubject variability. Since then, several studies have indicated that this adjustment of FEFs leads to clinically meaningful outcome variables. For example, lower FEF_{25-75}/FVC ratios were associated with a higher familial risk of developing COPD after smoking [6], with higher airway reactivity and sensitivity to methacholine [7, 8], and with higher airway reactivity to eucapnic hyperventilation with cold air [9].

In an editorial, THOMPSON [10] recommended a revival of the aforementioned dysanapsis concept. He pointed out that the adjusted FEFs lack normal reference values and are now subject to arbitrary cut-off values. Unfortunately, the 2012 Global Lung Initiative (GLI), in their capacity as a European Respiratory Society (ERS) task force, updated only the reference values for FEF_{75} and FEF_{25-75} , but refrained from providing reference equations for FVC-adjusted FEFs [11]. We agree with Thompson that reference values with lower and upper limits of normal may speed up the understanding, validation and implementation of the FEF/FVC outcomes.

In this study, we provide reference equations for the adjusted FEFs (FEF_{25}/FVC , FEF_{50}/FVC , FEF_{75}/FVC , and FEF_{25-75}/FVC). Furthermore, we provide an update of the unadjusted FEF equations and compare these with the equations from QUANJER *et al.* [12] and GLI [13].

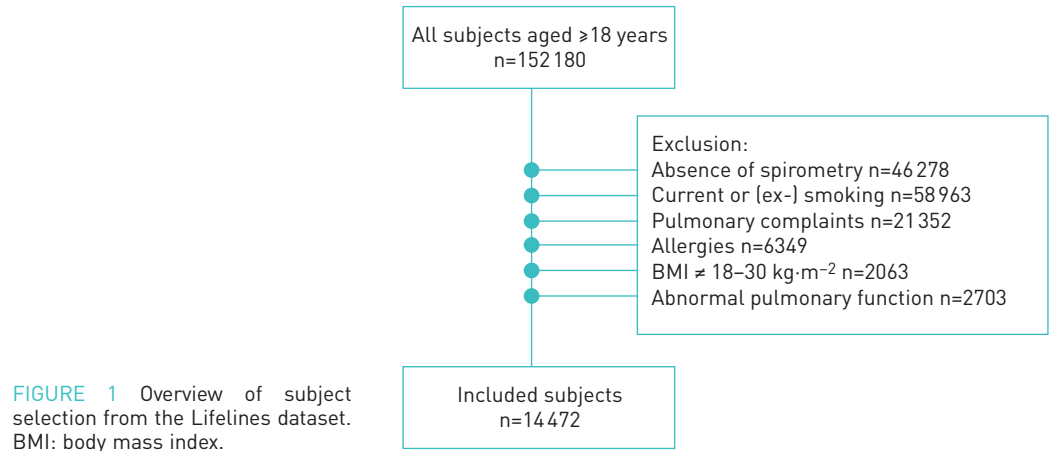
Methods

Subjects

Lifelines is a population-based prospective cohort (inclusion between 2003 and 2016), representative of the population of the Netherlands [14, 15], with an intended total follow-up of ≥ 30 years and a follow-up frequency of 5 years for measurements and 1.5 years for questionnaires. For this study, we used the baseline data of the 152 180 adult subjects who were aged ≥ 18 years and performed spirometry. We selected never-smokers, without any pulmonary complaints, who used no pulmonary medication, nor reported any allergies, and had a body mass index 18–30 $kg \cdot m^{-2}$. A further selection was performed based on normal pulmonary function (FEV_1 , FVC and FEV_1/FVC above the lower limit of normal [16]) and reliable spirometry, as judged by the pulmonary research technician obtaining the spirometry or a pulmonologist (figure 1 and supplementary material). For external validation we used 338 healthy never-smokers, without any past or present pulmonary complaints, and a normal pulmonary function. This was a combination of the unpublished Fiddle dataset approved by the ethical committee of the University Medical Center Groningen ($n=282$; see supplementary material) enriched with the healthy never-smokers of the NORM study ($n=56$) [17].

Data collection

The spirometry measurements were based on a full FVC manoeuvre performed according to the standardised operating procedure of the American Thoracic Society (ATS)/ERS task force [18]. In line with these guidelines, the best-effort set was used. All Lifelines data (including weight, height and



spirometry) are obtained according to standardised protocols, by trained technicians. Spirometry measurements were obtained on a PC-based SpiroPerfect with CardioPerfect software (Welch Allyn, Skaneateles Falls, NY, USA). The unpublished Fiddle dataset and the NORM study's dataset were obtained in the pulmonology department of the University Medical Center Groningen (Groningen, the Netherlands), using the MasterScreen PFT (Vyair Medical, Chicago, IL, USA).

Definition

The FVC-adjusted forced expiratory flow at 25%, 50%, 75% and mean 25%–75% of FVC were obtained by dividing FEF_{25} , FEF_{50} , FEF_{75} and FEF_{25-75} by the actually recorded FVC (equation 1). This is expressed in reciprocal time [19].

$$\text{adjusted } FEF_x(s^{-1}) = FEF_x(L \cdot s^{-1})/FVC(L) \quad (1)$$

Statistical analyses

To obtain reference equations for the adjusted and unadjusted FEFs, multiple linear regression was performed with explanatory variables age, weight and height. This regression was stratified by sex. All combinations of the explanatory variables were assessed, and the model with the highest adjusted R^2 was chosen as best model. Models were built on a random sample of 80% of the data (training set), using 10-fold cross-validation (R version 3.5.2, R-package: caret [20]). The obtained model was evaluated consecutively on the remaining 20% of the dataset (internal validation set) and on the Fiddle–NORM dataset (external validation set).

To check if the adjustment of the FEF decreased the variability of the FEFs, the coefficients of variation of the unadjusted and adjusted FEFs were compared using an asymptotic test (R-package: cvequality [21]).

To investigate to what extent the equations (our newly developed equations and the existing equations from QUANJER *et al.* [12] and GLI [13]) predict the unadjusted FEFs, the adjusted R^2 (explained variance) was used. The adjusted R^2 for the existing equations was calculated using R-package: rspirometry [22] and subsequently adjusted for the number of variables in the equation and the sample size of our dataset.

Results

A total of 14 472 healthy subjects were included from Lifelines; 6054 males and 8418 females (table 1 and supplementary table S2). The external validation set contained 338 subjects; 170 males and 168 females (supplementary table S2). The variability of the adjusted FEFs is depicted numerically in table 2 and visually in figure 2a and b. The coefficient of variation of the adjusted FEFs was significantly lower than that of the unadjusted FEFs, except for FEF_{25}/FVC (table 3).

Visual inspection of the data showed a fairly linear relationship between the adjusted FEFs and each of the explanatory variables, age, weight and height separately. Comparison of the eight possible (multiple) linear regression models showed that the best fit, defined as the highest adjusted R^2 , was obtained by including age, weight and height in all models for the adjusted FEFs, for both males and females (supplementary table S3). The best models for the adjusted FEFs (table 4) were internally and externally validated and showed numerically comparable fits in all datasets (table 4). Notably, the fit in the external validation set was better than in the training and internal validation set.

TABLE 1 Characteristics of groups used to train the reference equations

| | Male | Female |
|---|-----------|-----------|
| Subjects | 4846 | 6736 |
| Age[#] years | 42±11.6 | 42±12.4 |
| Height cm | 183±6.7 | 170±6.4 |
| Weight kg | 84±9.8 | 69±8.9 |
| BMI kg·m⁻² | 24.9±2.5 | 23.9±2.8 |
| FEV₁ L | 4.40±0.63 | 3.19±0.49 |
| FEV₁ % pred | 99±9.1 | 98±9.2 |
| FVC L | 5.58±0.74 | 3.99±0.56 |
| FVC % pred | 100±9.0 | 101±9.4 |
| FEV₁/FVC | 78.9±4.9 | 79.8±5.2 |
| FEF₂₅ L·s⁻¹ | 8.50±1.56 | 6.10±1.13 |
| FEF₂₅[¶] % pred | | |
| FEF₅₀ L·s⁻¹ | 4.87±1.22 | 3.67±0.91 |
| FEF₅₀[¶] % pred | | |
| FEF₇₅ L·s⁻¹ | 1.63±0.63 | 1.25±0.54 |
| FEF₇₅ % pred | 101±29.7 | 96±30.2 |
| FEF₂₅₋₇₅ L·s⁻¹ | 4.03±1.10 | 3.03±0.86 |
| FEF₂₅₋₇₅ % pred | 96±22.3 | 92±20.5 |

Data are presented as n or mean±sd. Predicted values are according to the Global Lung Initiative (GLI) equations [13]. BMI: body mass index; FEV₁: forced expiratory volume in 1 s; FVC: forced vital capacity; FEF_x: forced expiratory flow at 25%, 50%, 75% or 25–75% of FVC. #: age range (interquartile range) 18–80 (34–49) years and 18–85 (33–48) years for males and females, respectively; ¶: no reference equations available from the GLI. For characteristics of internal and external validation sets see supplementary table S2.

For the unadjusted FEFs, the best models were obtained when all explanatory variables were included (table 5). Internal and external validation showed numerically comparable fits (adjusted R²) in all datasets (table 5). These unadjusted FEF equations show similar fits on the training, internal and external validation sets compared to the 1993 equations from Quanjer and the GLI equations (table 6).

Discussion

In this study, we derived reference equations for the adjusted FEF values (FEF/FVC) and provided upper and lower limits of normal. Predicting the adjusted FEFs by using all tested explanatory variables (age, weight and height) resulted in the best reference equations, as based on the highest adjusted R². Furthermore, we showed that FVC adjustment of the FEF reduces the variability statistically. Finally, we calculated new reference equations for unadjusted FEF values and demonstrated similar predictive value compared to existing reference equations (presented by QUANJER *et al.* [12] in 1993 and the GLI task force in 2012 [13]).

We compared the coefficient of variance of the adjusted FEFs to those of the unadjusted FEFs and found that the adjustment significantly decreased the variability of this parameter (table 3), except in the FEF₂₅/FVC values. However, the magnitude of this reduction, was small and may therefore surpass clinical value. In 1974, GREEN *et al.* [5] adjusted flow (FEF) for the actual lung size (using vital capacity) in order to

TABLE 2 Overview of adjusted forced expiratory flow (FEF) for all groups

| | Male | Female |
|---|-----------|-----------|
| Subjects | 4846 | 6736 |
| FEF₂₅/FVC s⁻¹ | 1.54±0.30 | 1.54±0.29 |
| FEF₅₀/FVC s⁻¹ | 0.88±0.21 | 0.92±0.21 |
| FEF₇₅/FVC s⁻¹ | 0.29±0.11 | 0.31±0.12 |
| FEF₂₅₋₇₅/FVC s⁻¹ | 0.72±0.19 | 0.76±0.19 |

Data are presented as n or mean±sd. FEF_x: forced expiratory flow at 25%, 50%, 75% or 25–75% of forced vital capacity (FVC).

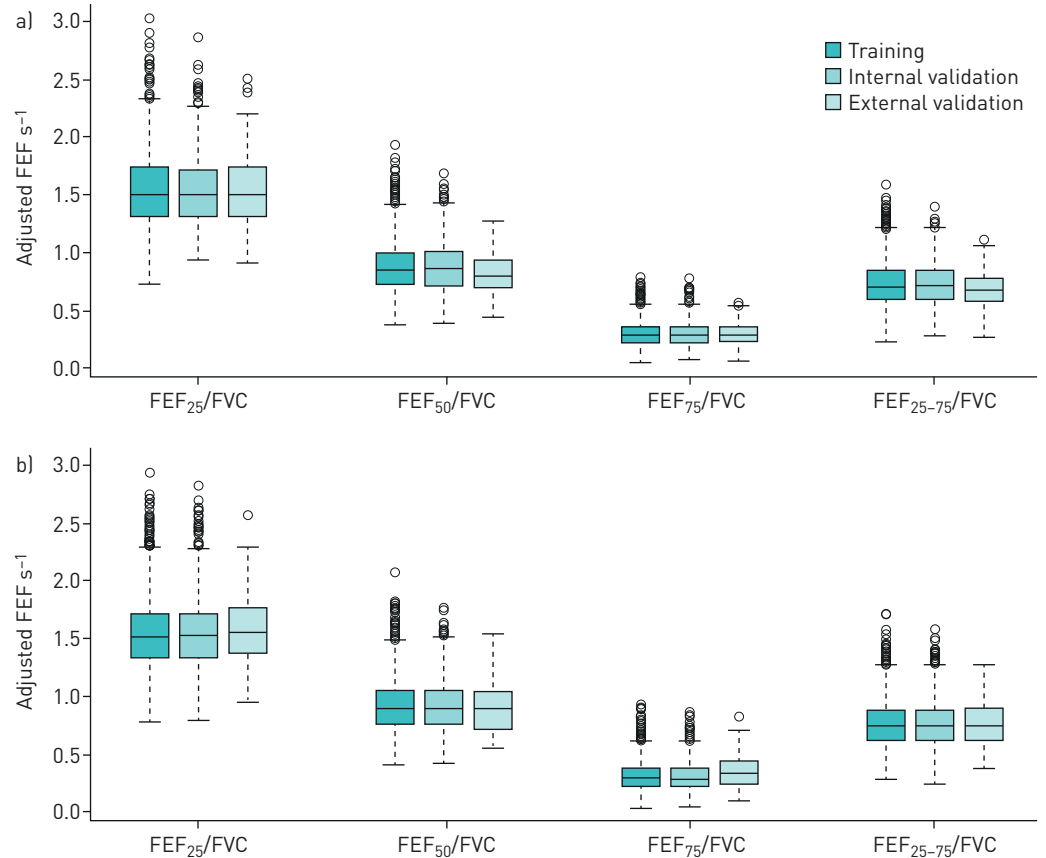


FIGURE 2 Variability in the adjusted forced expiratory flows (FEFs) for a) males and b) females, per datasets. FVC: forced vital capacity; FEF_x : FEF at 25%, 50%, 75% and 25–75% of FVC.

reduce the large intersubject variability. To their surprise, and in line with our findings, the variability only marginally decreased by this adjustment. As this result was nonsignificant in their small-sized study ($n=56$) they theorised that the FEFs are subject to substantial intersubject differences in airway size and function, independent of lung size (FVC). This independence is supported in later studies by showing that larger and central airway size is unrelated to lung size in normal adults, based on different imaging and functional techniques [23–25]. Regardless of the small reduction of the variability, we propose to use the adjusted FEFs. Their value is in the improved comparability between measurements due to less dependence on FVC performance.

| TABLE 3 Coefficient of variation for adjusted and unadjusted forced expiratory flows (FEFs) | | | |
|---|--|--|------------------------|
| | Adjusted FEF_x/FVC coefficient of variation % | Unadjusted FEF_x coefficient of variation % | p-value |
| Males | | | |
| FEF_{25} | 19.2 | 18.3 | 0.00027 |
| FEF_{50} | 23.9 | 25.1 | 0.00097 |
| FEF_{75} | 36.0 | 38.7 | 5.81×10^{-5} |
| FEF_{25-75} | 25.4 | 27.2 | 6.23×10^{-7} |
| Females | | | |
| FEF_{25} | 18.7 | 18.6 | 0.739 |
| FEF_{50} | 22.9 | 24.9 | 1.24×10^{-13} |
| FEF_{75} | 39.8 | 43.2 | 1.51×10^{-10} |
| FEF_{25-75} | 25.8 | 28.6 | 9.46×10^{-19} |

FEF_x : forced expiratory flow at 25%, 50%, 75% and 25–75% of forced vital capacity (FVC).

TABLE 4 Generated reference equations for adjusted forced expiratory flow (FEF) values

| | Equations | | Explained variance for the FEF/FVC equations % | | |
|---|-----------------------------|-------|--|---------------------|---------------------|
| | Reference equation | RSD | Training | Internal validation | External validation |
| Males | | | | | |
| FEF ₂₅ /FVC s ⁻¹ | 3.56+0.001A+0.003W-0.013H | 0.286 | 6.79 | 6.48 | 13.16 |
| FEF ₅₀ /FVC s ⁻¹ | 2.02-0.004A+0.001W-0.006H | 0.205 | 6.17 | 4.76 | 4.04 |
| FEF ₇₅ /FVC s ⁻¹ | 0.68-0.005A-0.0009W-0.0005H | 0.084 | 35.9 | 32.1 | 37.8 |
| FEF ₂₅₋₇₅ /FVC s ⁻¹ | 1.70-0.007A+0.0005W-0.004H | 0.169 | 16.6 | 13.8 | 14.9 |
| Females | | | | | |
| FEF ₂₅ /FVC s ⁻¹ | 3.52-0.0005A+0.003W-0.013H | 0.278 | 6.21 | 5.82 | 4.56 |
| FEF ₅₀ /FVC s ⁻¹ | 2.05-0.005A+0.001W-0.006H | 0.201 | 8.64 | 9.54 | 9.32 |
| FEF ₇₅ /FVC s ⁻¹ | 0.73-0.006A-0.001W-0.0004H | 0.095 | 41.0 | 41.5 | 46.0 |
| FEF ₂₅₋₇₅ /FVC s ⁻¹ | 1.76-0.008A+0.0009W-0.004H | 0.170 | 23.6 | 24.1 | 25.2 |

FVC: forced vital capacity; RSD: residual standard deviation; FEF_x: forced expiratory flow at 25%, 50%, 75% and 25-75% of FVC; A: age (years); W: weight (kg); H: height (cm). Lower limit of normal=value-1.64×RSD; upper limit of normal=value+1.64×RSD.

We checked the validity of the obtained reference equations on an internal and external dataset. The numerical comparison of the adjusted R² of the training set showed comparable adjusted R² values as compared to the internal and external validation set (table 4), demonstrating the validity of the equations in other populations. As the Lifelines dataset is a representative and generalisable sample of the Dutch population [15], we consider the obtained reference equations useful for the Dutch and comparable Caucasian populations.

Notwithstanding the similarity in the fit of the equations among the datasets, the reference equations of the adjusted FEFs had only weak to modest fits, expressed as the percentage of explained variance (adjusted R²). These were considerably lower than for the unadjusted values. We considered whether the overall relatively low predictive values of adjusted FEFs may be explained by the dependency of the FVC on age, weight, and height. *Via* the FVC adjustment, the FEFs are indirectly already adjusted for age, weight and height, as the FVC also depends on these. In other words, these explanatory variables theoretically lose explanatory value when introduced in a model that predicts the adjusted FEF. If FVC were completely dependent on age, weight and height, a prediction model for the adjusted FEF could even be independent of these explanatory variables, which would result in a fixed model. We therefore investigated whether a prediction model without explanatory variables would improve the explanatory value. In all cases, the model's adjusted R² including the explanatory variables was significantly higher than a model without explanatory variables (supplementary table S3). Hence, we conclude that a reference equation including age, weight and height is preferred over using one constant value as a reference for the adjusted FEFs.

TABLE 5 Generated reference equations unadjusted forced expiratory flow (FEF) values

| | Equations | | Explained variance for the unadjusted FEF equations % | | |
|--|----------------------------|-------|---|---------------------|---------------------|
| | Reference equation | RSD | Training | Internal validation | External validation |
| Males | | | | | |
| FEF ₂₅ L·s ⁻¹ | 0.50-0.021A+0.019W+0.040H | 1.478 | 9.73 | 12.1 | 11.6 |
| FEF ₅₀ L·s ⁻¹ | 0.14-0.038A+0.008W+0.031H | 1.099 | 19.5 | 19.8 | 20.7 |
| FEF ₇₅ L·s ⁻¹ | 0.14-0.034A-0.005W+0.018H | 0.469 | 46.8 | 44.7 | 55.0 |
| FEF ₂₅₋₇₅ L·s ⁻¹ | 0.37-0.048A+0.003W+0.029H | 0.903 | 32.6 | 31.9 | 39.2 |
| Females | | | | | |
| FEF ₂₅ L·s ⁻¹ | 1.21-0.026A+0.015W+0.029H | 1.039 | 15.6 | 17.0 | 15.4 |
| FEF ₅₀ L·s ⁻¹ | 0.51-0.033A+0.007W+0.024H | 0.778 | 27.1 | 27.7 | 29.8 |
| FEF ₇₅ L·s ⁻¹ | 0.35-0.028A-0.005W+0.014H | 0.375 | 51.4 | 52.0 | 56.8 |
| FEF ₂₅₋₇₅ L·s ⁻¹ | 0.74-0.041A+0.0015W+0.023H | 0.654 | 42.3 | 42.6 | 44.1 |

RSD: residual standard deviation; FEF_x: forced expiratory flow at 25%, 50%, 75% and 25-75% of forced vital capacity; A: age (years); W: weight (kg); H: height (cm). Lower limit of normal=value-1.64×RSD; upper limit of normal=value+1.64×RSD.

TABLE 6 Percentage of explained variance of existing reference equations on our datasets

| | Males | | | Females | | |
|--|----------|---------------------|---------------------|----------|---------------------|---------------------|
| | Training | Internal validation | External validation | Training | Internal validation | External validation |
| QUANJER [12] | | | | | | |
| FEF ₂₅ L·s ⁻¹ | 8.49 | 10.1 | 7.01 | 14.5 | 15.6 | 11.2 |
| FEF ₅₀ L·s ⁻¹ | 19.1 | 19.3 | 18.0 | 26.7 | 27.4 | 27.6 |
| FEF ₇₅ L·s ⁻¹ | 43.4 | 43.0 | 55.1 | 50.9 | 51.3 | 56.3 |
| FEF ₂₅₋₇₅ L·s ⁻¹ | 32.3 | 30.6 | 37.2 | 41.9 | 42.3 | 42.1 |
| GLI [13] | | | | | | |
| FEF ₇₅ L·s ⁻¹ | 47.4 | 46.3 | 55.1 | 53.0 | 53.6 | 58.3 |
| FEF ₂₅₋₇₅ L·s ⁻¹ | 32.8 | 32.0 | 38.2 | 42.5 | 42.7 | 42.9 |

Data are presented as %. FEF_x: forced expiratory flow at 25%, 50%, 75% and 25–75% of forced vital capacity.

Next to the overall predictive value it is striking that the explanatory value of the equations, for both the adjusted and unadjusted FEFs, increases towards the end of the expiration. This means that the (un)adjusted FEF₇₅ is more accurately predicted than the (un)adjusted FEF₂₅. This aligns with the theory that the end of the expiration is progressively effort independent [26]. At the beginning of the expiration, the (un)adjusted FEF depends more on explanatory variables unaccounted for in our reference models, like muscle strength or coordination. Towards the end of the expiration, factors less influenced by effort and practice, like age, weight and height, gain importance. Particularly age is an important contributing factor of the adjusted FEF₇₅ variability as the models without age had substantially lower adjusted R². This is in line with the well-known age-dependency of the unadjusted FEFs. We speculate that at older age, the small airways collapse more easily than at younger age, due to loss of retractile forces on the airways and loss of alveolar wall tension. In addition, decreased mucus clearance at higher age may contribute.

Even though the adjusted R² is comparable between datasets, the explanatory value of most of the obtained models was slightly higher in the external validation set. Additionally, the external dataset had smaller variability in adjusted FEF values (figure 1), which theoretically may be explained by the level of compliance to ERS/ATS spirometry acceptability criteria [18]: subjects selected for the external validation dataset had to be able to perform a spirometry completely according to these criteria whereas the Lifelines subjects needed to perform clinically reliable and reproducible spirometry. We therefore checked the variability and fit of the adjusted FEFs in the 1258 (out of our 14472) subjects able to perform spirometry completely compliant to the ERS/ATS criteria, and compared them with those of the external validation set (supplementary figure S4). However, the variability and fit of the adjusted FEFs from this Lifelines subset was not superior (supplementary table S5).

Next to adjusted FEF reference equations, we also generated unadjusted FEF reference equations (table 5). The Quanjer reference equations from 1993 were updated by the GLI task force in 2012, but only incorporated FEF₇₅ and FEF₂₅₋₇₅ equations and not FEF₂₅ and FEF₅₀. Compared to Quanjer, our equations showed only a minor improvement of the adjusted R² (table 6), probably due to a small cohort effect, indicating that current cohorts have a higher mean pulmonary function compared to older cohorts [27]. In fact, this cohort effect may be a reason to prefer our equations over the unadjusted FEF equations of Quanjer. To our surprise, the equations presented by the GLI, which were obtained with advanced statistical techniques on a more recent and healthier cohort (compared to Quanjer), and which incorporate an age-spline [13], only had slightly higher adjusted R² values compared to both our equations and the 1993 equations. Apparently, the log-transformation of the explanatory variables age and height, and the age-spline have limited additional value. The need for the age-spline in the dataset used by GLI may originate from the right-skewed age distribution of the sample, with 47% of the subjects aged <20 years [13]. In contrast, our Lifelines data was normally distributed around age 42 years and the Fiddle dataset was uniformly distributed with regards to age. This may explain the comparability of our linear reference equations to the GLI equations and it supports our choice to keep the model simple, without the introduction of a spline.

The strength of this study is the large sample size, collected according to the same protocol, using a unified set up by technicians with equivalent training. This ensured a large measurement homogeneity and through the selection process the data was generalisable for the Caucasian population [15]. Furthermore, all spirometries with questionable reproducibility or ERS/ATS compliance were assessed by an independent pulmonologist. In comparison, the GLI task force had a larger, but more heterogeneous

dataset, as it consists of a combination of databases from 72 studies from 33 countries [13] and was therefore potentially measured according to different protocols, with several spirometry devices, operated by differently trained pulmonary technicians. Furthermore, combining different databases is likely to introduce differences in data quality. The dataset used by GLI covers a broader spectrum of subjects, which aids generalisability at the cost of introducing a larger variability in the dataset, which subsequently introduces the need for the more complicated statistics.

Conclusion

Using >14000 healthy subjects, we developed reference equations for FVC-adjusted FEF values. We demonstrate acceptable fits, both in internal and external datasets. Additionally, we demonstrate minimally, yet significantly improved, variability as compared to unadjusted FEF values. A next step will be to evaluate the clinical relevance of the obtained reference equations in subjects with established airway disease.

Acknowledgements: The authors wish to thank all participants in Lifelines, the NORM study and the department's internal dataset. Furthermore, we wish to thank emeritus professor Dirkje S. Postma (University of Groningen, University Medical Centre Groningen, Groningen, the Netherlands) for her contributions to the quality checks of the Lifelines spirometry data. Last, we would like to thank the pulmonary function technicians of Lifelines and our own lab, in particular Martijn P. Farenhorst and Cindy C.S. Alberts-Poots (University of Groningen, University Medical Centre Groningen) as they registered all participants and performed the data collection for the Fiddle dataset.

Conflict of interest: C.A. Cox has nothing to disclose. J.M. Vonk has nothing to disclose. H.A.M. Kerstjens reports research grants from GSK, Novartis and Boehringer, and fees for consultancy on advisory boards from GSK, Novartis and Boehringer, all paid to his institution. M. van den Berge reports grants paid to his university from AstraZeneca, TEVA, GSK and Chiesi outside the submitted work. N.H.T. ten Hacken has nothing to disclose.

Support statement: This study was supported by Universitair Medisch Centrum Groningen, Pieken in de Delta, Ministerie van Economische Zaken, the European Fund for Regional Development, Nederlandse Overheid, Nederlandse Organisatie voor Wetenschappelijk Onderzoek, Samenwerkingsverband Noord-Nederland, Koninklijke Nederlandse Akademie van Wetenschappen, Rijksuniversiteit Groningen, the Target Corporation, Provincie Groningen, BBMRI-NL and Provincie Drente. Funding information for this article has been deposited with the Crossref Funder Registry.

References

- Gibson GJ. Spirometry: then and now. *Breathe* 2005; 1: 206–216.
- Cochrane GM, Prieto F, Clark TJH. Intrasubject variability of maximal expiratory flow volume curve. *Thorax* 1977; 32: 171–176.
- Pellegrino R, Viegi G, Brusasco V, et al. Interpretative strategies for lung function tests. *Eur Respir J* 2005; 26: 948–968.
- Miller MR, Crapo R, Hankinson J, et al. General considerations for lung function testing. *Eur Respir J* 2005; 26: 153–161.
- Green M, Mead J, Turner MJ. Variability of maximum expiratory flow-volume curves. *J Appl Physiol* 1974; 37: 67–74.
- DeMeo DL, Carey VJ, Chapman HA, et al. Familial aggregation of FEF_{25–75} and FEF_{25–75}/FVC in families with severe, early onset COPD. *Thorax* 2004; 59: 396–400.
- Parker AL, Abu-Hijleh M, McCool FD. Ratio between forced expiratory flow between 25% and 75% of vital capacity and FVC is a determinant of airway reactivity and sensitivity to methacholine. *Chest* 2003; 124: 63–69.
- Litonjua AA, Sparrow D, Weiss ST. The FEF_{25–75}/FVC ratio is associated with methacholine airway responsiveness: the normative aging study. *Am J Respir Crit Care Med* 1999; 159: 1574–1579.
- Tager IB, Weiss ST, Muñoz A, et al. Determinants of response to eucapnic hyperventilation with cold air in a population-based study. *Am Rev Respir Dis* 1986; 134: 502–508.
- Thompson BR. Dysanapsis – once believed to be a physiological curiosity – is now clinically important. *Am J Respir Crit Care Med* 2017; 195: 277–278.
- Quanjer PH, Stocks J, Cole TJ. 2012. GLI-2012 All-Age Multi-Ethnic Reference Values for Spirometry; Advantages Consequences; Interpretation of Spirometric Data available from: <https://www.ers-education.org/lrmedia/2012/pdf/266696.pdf>
- Quanjer PH, Tammeling GJ, Cotes JE, et al. Lung volumes and forced ventilatory flows Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J* 1993; 6: Suppl. 16, 5–40.
- Quanjer PH, Stanojevic S, Cole TJ, et al. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.
- Scholtens S, Smidt N, Swertz MA, et al. Cohort profile: LifeLines, a three-generation cohort study and biobank. *Int J Epidemiol* 2015; 44: 1172–1180.
- Klijs B, Scholtens S, Mandemakers JJ, et al. Representativeness of the LifeLines cohort study. *PLoS One* 2015; 10: e0137203.
- Quanjer PH, Stanojevic S, Cole TJ, et al. 2015. Implementing GLI-2012 Regression Equations. Global Lung Function Initiative. <https://www.ers-education.org/lrMedia/2013/pdf/266709.pdf>
- Cox CA, Vonk JM, Kerstjens HAM, et al. Factors associated with hyperresponsiveness to adenosine 5'-monophosphate in healthy subjects. *Allergy* 2019; 74: 2268–2270.
- Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.
- Mead J. Dysanapsis in normal lungs assessed by the relationship between maximal flow, static recoil, and vital capacity. *Am Rev Respir Dis* 1980; 121: 339–342.

- 20 Kuhn M. Building predictive models in R using the caret package. *J Stat Softw* 2008; 28: 1–26.
- 21 Marwick B, Krishnamoorthy K. 2019. cvequality: Tests for the Equality of Coefficients of Variation from Multiple Groups. <https://github.com/benmarwick/cvequality>
- 22 Lytras T. 2017. rspiro: Implementation of the GLLI-2012 Equations in R. <https://github.com/thlytras/rspiro>
- 23 Martin TR, Castile RG, Fredberg JJ, *et al.* Airway size is related to sex but not lung size in normal adults. *J Appl Physiol* 1987; 63: 2042–2047.
- 24 Sheel AW, Guenette JA, Yuan R, *et al.* Evidence for dysanapsis using computed tomographic imaging of the airways in older ex-smokers. *J Appl Physiol* 2009; 107: 1622–1628.
- 25 Diaz AA, Rahaghi FN, Ross JC, *et al.* Understanding the contribution of native tracheobronchial structure to lung function: CT assessment of airway morphology in never smokers. *Respir Res* 2015; 16: 23.
- 26 Morris JF. Spirometry in the evaluation of pulmonary function. *West J Med* 1976; 125: 110–118.
- 27 Kerstjens HAM, Rijcken B, Schouten JP, *et al.* Decline of FEV₁ by age and smoking status: facts, figures, and fallacies. *Thorax* 1997; 52: 820–827.