DATA

Canadian CF Registry (CCFR)

The Canadian CF Registry (CCFR) contains detailed demographic and annual clinical information on CF patients receiving clinical care at 42 accredited CF centres across Canada. All patients provided consent for their data to be collected in the CCFR and used for research purposes. CCFR offers a comprehensive picture of the CF population, with more than 95% of Canadians living with CF being captured in the registry.[1] Data errors and inconsistency are systematically resolved by routine data validation procedures and by cross-referencing with original sources at the reporting CF centres.

US CF-specific reference equations

To compare with the US, we obtained the most up-to-date US reference equations from www.karlin.mff.cuni.cz/~kulich, constructed based on the US Cystic Fibrosis Foundation (CFF) registry data collected from January 1, 2001 to December 31, 2006.

European CF-specific reference equations

For the European countries, the reference equations were defined based on a combination of data from the French CF Modifier Gene Study (MUCONAT) and European Cystic Fibrosis Society Patient Registry (ECFSPR). The European reference equations grouped data from France and several other European countries, collected respectively, from 2008 to 2010 and from 2004 to 2007.[2] The European reference equations were not available publicly, we therefore extracted the approximate coordinates from the published figures[2] using Plot Digitizer V2.6.8 then applied smoothing. To assess the reliability of this approach, we also implemented this software on the graphs for the US CF-specific percentiles[3] and compared the estimated coordinates to the published ones. In the following figure S8, the reproduced reference percentiles are compared to the actual percentiles from the publication, showing both percentiles are nearly identical:

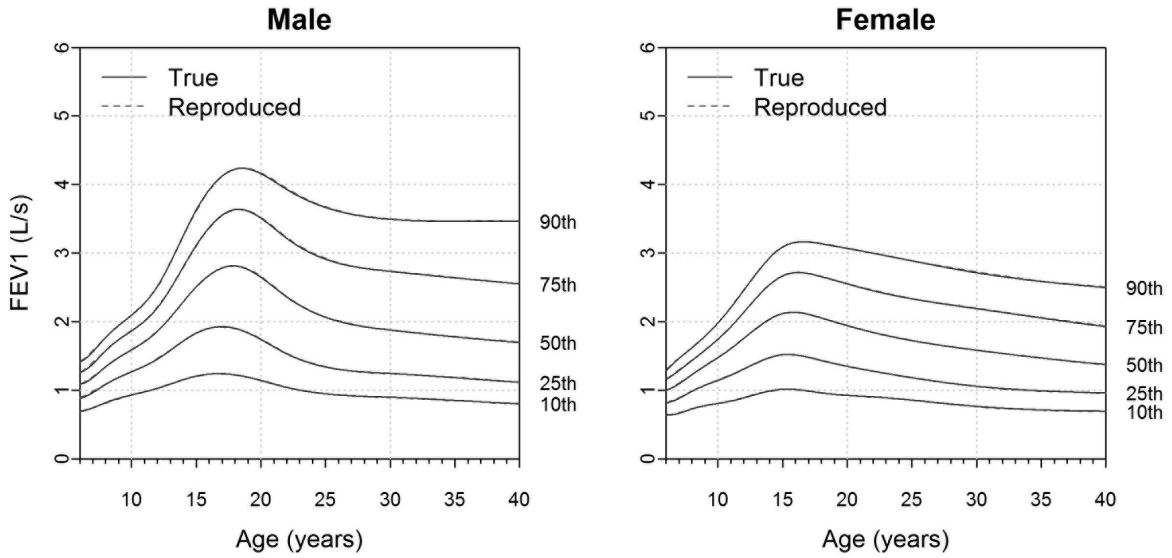

Figure S8. Sensitivity analysis of using the Plot Digitizer V2.6.8 (dashed line) for calculation of the US CF-specific FEV1 percentiles adjusted for age, stratified by sex; the actual percentiles from the publication are displayed using solid lines.

Data inclusion and exclusion criteria

Height and lung function measurements from the first stable measurement of each year were used for these analyses. A stable measurement was defined as one that is taken from a routine outpatient clinic visit when the patient was not being treated for a pulmonary exacerbation. Due to the possibility of seasonality bias, we also calculated the percentiles using a random selection of annual measurements. The resulting FEV1 percentiles were extremely similar so we concluded that using the first stable measurement was a robust approach. Patients under the age of 6 and over 50 years of age were excluded. For inter-country or temporal comparisons, this age limit was changed to 40 years to match previous study criteria.[2-3] Only height measurements between 105cm and 190cm, and between 105cm and 180cm, were included for male and female patients, respectively. Since clinical outcomes post-transplant, particularly FEV1, do not represent CF lung disease, FEV1 measurements post-transplant were excluded. The data inclusion/exclusion for the Canadian contemporary CF reference equations are shown in supplementary figure S1.

To develop contemporary Canadian CF-specific reference equations for FEV1, we used the subset of data collected from January 1, 2008 to December 31, 2014. These contemporary percentiles were compared to CCFR data from January 1, 2000 to December 31, 2007 to evaluate changes over time within Canada.

For each inter-country comparison, we used the same data collection period, data inclusion/exclusion criteria and statistical method as reported in the original articles.

STATISTICAL METHODS

Regression model

We used quantile regression[4] to estimate each Canadian CF-specific FEV1%ile adjusting for age and/or height, separately by sex. Quantile regression is a procedure that has been implemented for spirometry in other populations.[2-3, 5-6] To account for the non-linearity in the predictors, we used an additive model and applied smoothing using cubic B-spline bases, consisting of 6 interior knots for each predictor.[3] Since each percentile is estimated separately, percentile curves can intersect. As a remedial measure, we applied non-crossing constraints.[7] Such constraints were removed for the international comparisons, so as to standardize the statistical methods used across the different populations studied. Due to a low frequency of non-Caucasians (<5%) in the CCFR, we did not adjust for race in our regression model as the low sample size would yield poor estimates of FEV1%iles. Sensitivity analysis displayed almost identical curves when comparing percentiles for the full Canadian CF population versus those restricted to Caucasians.

The model specification for the contemporary Canadian reference equations is shown as follows:

$$Q_\tau(\text{FEV1}|\text{Age},\text{Height}) = \beta_0(\tau) + \sum_{k_1=1}^{K_1} B_{k_1}(\text{Age})\beta_{k_1}(\tau) + \sum_{k_2=1}^{K_2} B_{k_2}(\text{Height})\beta_{k_2}(\tau)$$

constrained to:

$$\begin{cases} Q_{0.50} < Q_{0.51} < \cdots < Q_{0.99} \\ Q_{0.50} > Q_{0.49} > \cdots > Q_{0.01} \end{cases}$$

where the range of age is [6, 50], the range of height is [105, 190] for males and [105, 180] for females, $Q_\tau(\cdot)$ is the τ's conditional quantile ($\tau = 0.01, \ldots, 0.99$), $B_k(x)$ is the k^{th} B-spline basis function ($k = 1, \ldots, K$) for a covariate x and $\beta_k(\tau)$ is the τ^{th} quantile-specific regression coefficients corresponding to k^{th} B-spline basis function. To generate the cubic B-spline basis functions for age, the boundary knots were placed at 6 and 50 years and 6 interior knots were at 8, 11, 15, 20, 25 and 35 years, which led to 9 basis

functions ($K_1 = 9$). For height, the boundary knots were 105 and 190cm for males and 105 and 180cm for females. 6 interior knots were placed at 115, 125, 135, 150, 160 and 170cm and for females, at 115, 125, 135, 145, 155 and 165cm, leading to 9 basis functions ($K_2 = 9$).

The non-crossing constraints were not implemented in the US and European CF-specific FEV1%ile studies [2-3]. We compared the constrained FEV1 percentiles to the unconstrained percentiles and observed that departure between the two approaches was limited to the upper percentiles (90-99th) for females at older ages and heights. This was due to more sparse data in females (greater mortality) at older ages especially for the higher percentiles.

Statistical tests

To estimate the variance of the conditional quantiles to enable statistical testing, we accounted for the correlation of FEV1 measurements within individuals by using a cluster bootstrap with 1000 replicates.[8] In this procedure, patients were sampled first, retaining the selected patient's FEV1 measurements, preserving the correlated structure of their FEV1. To test the difference in age or height-adjusted CF-specific FEV1%iles between Canada and the US, between Canada and Europe or between two distinct time periods, we used a bootstrap test for the difference in the area under the curve (AUC) with 1000 replicates at 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles. For each comparison, Bonferroni adjustment was used to account for the 5 statistical tests.

We treated the estimated FEV1%iles for the US and European data as true values in the bootstrap tests since the raw data were not available, consequently underestimating the true variance of the estimated FEV1%iles. However, we expect the variance to be much smaller for the US and European samples than for Canada because the number of measurements was far greater. In the US study, 287,108 measurements from 28,000 unique patients were included compared to 13,942 measurements from 3,156 unique patients from Canada during the same time period, over a 20-fold increase in the number of measurements. The European study included 50,482 measurements from 16,781 unique patients compared to 17,096 measurements from 3,338 unique patients for Canada, approximately 3 times greater than our data.

In all comparisons of FEV1%iles, we used the difference in the area under the curve (AUC) as a measure of a general improvement of lung function at the specific percentile; this procedure was also implemented in [2]. It represents the cumulative lung function over the range of age or height for a given percentile. We note that, although significance testing of AUCs is useful to assess the overall improvements of lung function, it does not demonstrate where one percentile curve is significantly different from another.

We calculated 99% pairwise confidence bands to determine the range of age or height where the FEV1%ile of one group is significantly higher than another. For temporal comparison of age-adjusted FEV1%iles, significant differences were observed for males at 15.8 - 18.3 and 19.6 - 21.2 years for the 10th percentile, 19.9 - 28.9 years for 25th percentile, 16.5 - 31.4 years for the 50th percentile, 23.5 - 34.1 years for the 75th percentile and 32.5 - 33.8 years for the 90th percentile. For females, the difference was significant at 32.6 - 33.6 years for the 10th percentile and 31.2 - 33.6 years for the 25th percentile.
For the comparison between US and Canadian males, significant differences were observed at 7 - 18.1 and 25.7 - 39.7 years for the 10th percentile, 7.6 - 22.5 and 24.6 - 39.7 years for the 25th percentile, 7.6 - 40 years for the 50th percentile, 7.5 - 18.1 and 37.8 - 39.6 years for the 75th percentile and 9.2 - 13.8 and 15 - 16.9 for the 90th percentile. For females, the differences were significant at 6.5 - 32.4 years for the 10th percentile, 6.6 - 29.3 and 38.8 - 40 years for the 25th percentile, 6 - 27.4 and 38.1 - 40 years for the 50th percentile, 7.5 - 16.5 and 19 - 24.3 years for the 75th percentile and 6.9 - 8, 11.5 - 16.4 and 38.9 - 39.4 years for the 90th percentile.

For the comparison between European and Canadian males, the differences were significant at 13.3 - 39 years for the 10th percentile, 13 - 40 years for the 25th percentile, 12.4 - 17.2, 19.5 - 28.7 and 33.4 - 40 years for the 50th percentile and 6 - 9.4, 13 - 16.8 and 21.6 - 25.9 years for the 75th percentile. For females, the corresponding age range is 11.7 - 32.7 and 38.4 - 40 years for the 10th percentile, 9.9 - 15.9, 21.6 - 31.5 and 38.2 - 40 years for the 25th percentile, 6 - 7.3, 11.6 - 14.3, 21.1 - 30.4 and 38.7 - 40 years for the 50th percentile, 37.6 - 40 years for the 75th percentile and 38.8 - 40 years for the 90th percentile.

Assessing Generalized Additive Models for Location, Scale and Shape (GAMLSS) as an alternative model for calculating reference equations
As an alternative to quantile regression, we considered Generalized Additive Models for Location, Scale and Shape (GAMLSS).[9] Although previous CF-specific FEV1%iles were calculated using quantile regression,[2-3] GAMLSS has been used to calculate standard reference equations compared to healthy populations.[10-11] GAMLSS assumes a distribution for the response variable, allowing one to model not only the mean but also the variance, skewness and kurtosis using smoothing functions.[9] We used the Box-Cox Power Exponential distribution for GAMLSS and employed the same smoothing technique as in our quantile regression model, outlined above. Percentile curves were compared graphically.

All analyses were carried out using R V.3.3.2.[12] We used the packages "quantreg" and "gamlss" for the quantile regression and GAMLSS, respectively.

REFERENCES

1 Sykes J, Stanojevic S, Stephenson AL, *et al*. A standardized approach to estimating survival statistics for population-based cystic fibrosis registry cohorts. *J Clin Epidemiol* 2016;70:206–213.

2 Boëlle PY, Viviani L, Corvol H, *et al*. Reference percentiles for FEV1 and BMI in European children and adults with cystic fibrosis. *Orphanet J Rare Dis* 2012;7:64.

3 Kulich M, Rosenfeld M, Ramsey B, *et al*. Disease-specific reference equations for lung function in patients with cystic fibrosis. *Am J Respir Crit Care Med* 2005;172:885–891

4 Koenker R, Bassett G. Regression quantiles. *Econometrica* 1978;46:33–50.

5 Wei Y, Pere A, He X, *et al*. Quantile regression methods for reference growth charts, *Stat Med* 2006;25:1369-1382

6 Rochat MK, Laubender RP, Wildhaber J, *et al*. Spirometry reference equations for central European populations from school age to old age. *PLoS One* 2013;8:e52619.

7 Bondell HD, Reich BJ, Wang H. Non-crossing quantile regression curve estimation. *Biometrika* 2010;97:825–838

8 Sherman M, leCessie S. A comparison between bootstrap methods and generalized estimating equations for correlated outcomes in generalized linear models. *Commun Stat Simul Comput* 1997;26:901-25.

9 Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape (with discussion). *Appl Statist* 2005;54:507–554.

10 Stanojevic S, Wade A, Cole TJ, *et al*. Reference ranges for spirometry across all ages: a new approach. *Am J Respir Crit Care Med* 2008;177:253–260.

11 Quanjer PH, Stanojevic S, Stocks J, *et al*. Multi-ethnic reference values for spirometry for the 3–95 year age range: the global lung function 2012 equations. *Eur Respir J* 2012;40:1324-1343.

12 R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, 2013.

13 World Health Organization. BMI classification. 2006. Accessed at http://apps.who.int/bmi/index.jsp?introPage=intro_3.html on 1 May 2017.