**DOES CURRENT REPORTING OF LUNG FUNCTION BY THE UK CYSTIC FIBROSIS REGISTRY ALLOW A FAIR COMPARISON OF ADULT CENTRES?**

**Dr Julia Anne Nightingalea and Professor Clive Osmondb**

Running title: A comparison of UK adult CF centre respiratory outcomes.

Corresponding author:

Dr Julia Anne Nightingale

aDepartment of Adult Cystic Fibrosis

University Hospital Southampton NHS Trust

Southampton General Hospital

Tremona Road

Southampton

SO16 6YD

United Kingdom

[julia.nightingale@uhs.nhs.uk](mailto:julia.nightingale@uhs.nhs.uk)

Co-author:

Professor Clive Osmond.

Professor of Biostatistics

bMRC Lifecourse Epidemiology Unit

University of Southampton

Southampton

SO16 6YD

United Kingdom

[co@mrc.soton.ac.uk](mailto:co@mrc.soton.ac.uk)

Key words:

Cystic Fibrosis

Clinical Epidemiology

Respiratory Measurement

**Abstract**

*Background***:** Outcome data for UK cystic fibrosis centres are publicly available in an annual report, which ranks centres by median FEV1% predicted. We wished to assess whether there are differences in lung function outcomes between adult centres that might imply differing standards of care.

*Methods:*UKRegistry data from 4,761 subjects at 34 anonymised adult centres were used to calculate mean FEV1% and rate of change of lung function for 2007-13. These measures were used to rank centres and compare outcomes.

*Results:* There are minor differences between centres for mean FEV1% for some years of the study and for rate of change of lung function over the study period. However, rankings are critically dependent on the outcome measure chosen and centre variation becomes negligible once patient population characteristics are taken into account.

*Conclusions:*We have demonstrated that the ranking of centres is biased and any apparent difference in respiratory outcomes is unlikely to be related to differing standards of care between centres.

**1. Introduction**

Cystic fibrosis (CF) registries are powerful tools to drive improvements in patient care [1]. Registry data have been used to make comparisons between centres both nationally [2,3] and internationally [4,5]. These comparisons have been used to benchmark services and further the understanding of which treatment strategies may be useful in improving clinical outcomes for patients [2,3,6-8].

The United Kingdom (UK) CF Registry collects anonymised clinical data annually for all consenting CF patients cared for in specialist centres within the UK. Data are publicly available in an annual report with comparison between centres for both adult and paediatric clinics. The report includes a respiratory outcome measure of median FEV1% predicted (FEV1%) for each centre presented as a ranked graph and also in the most recent report as funnel plots [9]. A funnel plot is constructed by plotting the measured outcome against centre size, so that control limits form a funnel around the median value [10]. This allows a visual representation of values taking into account expected natural variation and centres that lie outside of this expected variation might have values that reflect a real difference from the average. These representations may be taken to imply that centres ranked higher for median FEV1% are ‘better’ than those with a lower ranking and those lying outside the control limits of a funnel plot are ‘better’ or ‘worse’ than expected.

One of the stated aims of the Registry report is to ‘improve the delivery of care’ and many centres (and possibly patients) use data informally to assess their outcomes in relation to other centres. As such, it is important that we understand both the significance and limitations of the results presented. We wished to determine whether the rankings published in the UK Registry report are indicative of standards of care within adult CF units. In an attempt to answer this query we addressed the following points. Firstly, whether it is possible to demonstrate a statistically significant difference between centres for FEV1%. Secondly, to assess the stability of rankings both with time and also with the use of an alternative respiratory outcome measure to rank centres. Finally, to explore how results are influenced by the characteristics of the population attending a centre.

**2. Methods**

*2.1 Study population*

The study population included all patients with a diagnosis of CF enrolled in the UK CF Registry in the years 2007-2013. National Health Service (NHS) research ethics approval was granted for the Registry and each patient or legally authorised representative provided written informed consent for data collection and research. Under the terms of the NHS ethics approval, the UK CF Trust steering committee approved this study. The Registry provided anonymised annual review data with one value for each patient per year for the seven years 2007-13 [11]. From 2012 onwards, the registry stipulated that the recorded data be the best in year value; prior to this date, values represent those recorded at annual review. Data points were included if there was a measurement of FEV1, height and age and subjects had remained in one centre throughout the study period. We also excluded data for those aged below 16 years or above 70 years, those whose gender and date of birth were not reported consistently across the years, and those whose FEV1 exceeded 155% predicted. We excluded centres for which there were fewer than 40 reports over the study period. 9.9% of the original records were excluded; two thirds because the subjects had changed centres within the study period. We studied the remaining 23,433 reports on 4,761 subjects treated at one of 34 anonymised adult centres All subjects were categorised for genotype (ΔF508 homozygous or other genotype) and for pancreatic enzyme use (yes, no or unknown).

*2.2 Comparison of centres*

We represented each FEV1 measurement as the percentage of the age-, sex-, height- and ethnicity-adjusted FEV1 value predicted for a healthy population using the 2012 global lung function equations [12]. Centres were compared for mean FEV1% for each year of the study using a linear mixed model.

*2.3 Lung function rate of change*

For each subject we plotted the FEV1% against time and used the regression slope as a summary measure of rate of change of lung function. In more effective centres subjects’ lung function would be expected to decline more slowly, allowing comparison of centres. These regression slopes are less reliable for subjects with fewer annual reports than for those with all seven, so we controlled for this in the analysis by using a Fisher-Yates transformation for each number of reports [13]. This led to a set of slopes for the 4761 subjects that had mean zero and standard deviation unity. We used the mean value of these slopes to compare rate of change in FEV1% across centres.

*2.4 Ranking of centres*

The 2013 dataset was used to calculate mean FEV1% predicted for each centre and these values were used to rank centres. We calculated Pearson correlations between rankings for each year of the study to explore how centre rankings varied with time. We also ranked centres using median FEV1% predicted and rate of change of FEV1% as alternative outcome measures.

*2.5 Controlling for patient characteristics within centres*

We used a mixed linear regression model to estimate the associations of age, sex, genotype and use of pancreatic enzymes on our lung function outcome measures and to also measure the two components of variance: between centre and between subjects within a centre. SPSS version 24 (IBM Corporation) was used for analysis.

**3. Results**

*3.1 Study population*

We studied 23,433 reports on 4,761 subjects treated at one of 34 anonymised adult centres. 574 subjects had complete data for two years within the study time period; 562 had data for three years; 756 for four; 715 for five; 1078 for six; and 1076 had data from all seven years. In 2007 there were 1956 reports, rising to 4032 in 2011, subsequently, report numbers remained relatively stable (supplementary table).

Mean FEV1% predicted for 2013 was used to rank centres (numbered 1 to 33). One centre did not contribute any data for 2013 and so was excluded from further analysis. Demographic details for each centre are shown in Table 1.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Centre ranked for mean FEV1%  pred | Number  of  patients | Mean  FEV1%  pred 2013 | Mean  age  (years) | Percent  women | Percent\*  ΔF508 homozygous | Percent\*  taking  pancreatic  enzymes |
|  |  |  |  |  |  |  |
| 1 | 25 | 72.3 | 25.4 | 36.0 | 36.0 | 76.0 |
| 2 | 35 | 68.5 | 30.5 | 34.3 | 57.1 | 82.9 |
| 3 | 122 | 68.3 | 28.2 | 48.4 | 50.4 | 81.0 |
| 4 | 173 | 66.3 | 32.1 | 43.4 | 40.0 | 73.9 |
| 5 | 227 | 65.9 | 30.2 | 45.4 | 51.1 | 85.0 |
| 6 | 33 | 65.3 | 31.5 | 54.5 | 48.5 | 78.1 |
| 7 | 12 | 65.1 | 32.7 | 41.7 | 33.3 | 90.9 |
| 8 | 69 | 64.8 | 29.1 | 50.7 | 48.5 | 75.0 |
| 9 | 15 | 64.6 | 31.9 | 26.7 | 33.3 | 66.7 |
| 10 | 122 | 63.7 | 30.8 | 48.4 | 36.9 | 80.8 |
| 11 | 110 | 63.2 | 29.3 | 46.4 | 31.4 | 83.0 |
| 12 | 22 | 63.2 | 24.9 | 59.1 | 63.6 | 95.2 |
| 13 | 62 | 63.1 | 31.3 | 40.3 | 51.6 | 83.1 |
| 14 | 142 | 63.0 | 30.3 | 51.4 | 45.3 | 76.3 |
| 15 | 261 | 62.9 | 31.4 | 39.8 | 51.5 | 89.6 |
| 16 | 208 | 62.8 | 30.8 | 43.8 | 49.0 | 86.6 |
| 17 | 99 | 62.7 | 31.0 | 43.4 | 57.7 | 88.8 |
| 18 | 203 | 62.5 | 29.4 | 47.8 | 52.2 | 88.7 |
| 19 | 127 | 62.4 | 26.9 | 49.6 | 55.1 | 91.8 |
| 20 | 321 | 62.1 | 30.5 | 41.4 | 52.3 | 87.9 |
| 21 | 186 | 61.7 | 29.4 | 43.5 | 43.5 | 82.8 |
| 22 | 39 | 61.5 | 27.8 | 48.7 | 64.1 | 94.7 |
| 23 | 341 | 60.5 | 31.4 | 43.1 | 58.2 | 88.0 |
| 24 | 74 | 60.4 | 32.8 | 45.9 | 54.1 | 87.0 |
| 25 | 92 | 60.4 | 29.4 | 44.6 | 48.4 | 91.7 |
| 26 | 556 | 60.3 | 33.8 | 45.7 | 48.5 | 83.3 |
| 27 | 26 | 58.9 | 30.5 | 50.0 | 50.0 | 87.0 |
| 28 | 18 | 58.0 | 30.6 | 61.1 | 55.6 | 88.2 |
| 29 | 42 | 57.1 | 31.4 | 47.6 | 57.1 | 81.1 |
| 30 | 50 | 55.7 | 27.2 | 50.0 | 53.1 | 86.7 |
| 31 | 47 | 55.1 | 30.3 | 42.6 | 44.7 | 87.2 |
| 32 | 42 | 54.8 | 27.8 | 61.9 | 48.8 | 87.5 |
| 33 | 26 | 52.0 | 29.1 | 34.6 | 69.2 | 96.0 |
|  |  |  |  |  |  |  |
| All patients | 3927 | 62.3 | 30.7 | 45.1 | 49.8 | 85.1 |
|  |  |  |  |  |  |  |
| P | - | 0.22 | 0.01 | 0.6 | <0.001 | <0.001 |
|  |  |  |  |  |  |  |

P measures the extent of between-centre variability.

\* Unknown values excluded.

**Table 1** Demographic details of patients by centre in 2013 (ranked for mean FEV1% predicted)

*3.2 Comparison of centres*

There was no significant difference in mean FEV1% predicted between centres in 2013. However, there were marginal statistically significant differences for some years (Table 2). In all years the degree of variation in values derives much more from differing populations within a clinic than from variation between clinics. There was a highly significant difference in the number of patients who were homozygous for the ΔF508 gene and the number of patients using pancreatic enzymes between centres and also a measurable difference in the mean age (Table 1).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Year | FEV1 % predicted | | | |
|
| Within centre SD | Between centre SD | p-value | ICC (%) |
|
|  |  |  |  |  |
| 2007 | 23.3 | 2.2 | 0.11 | 0.9 |
| 2008 | 23.3 | 2.1 | 0.13 | 0.8 |
| 2009 | 23.4 | 2.8 | 0.03 | 1.4 |
| 2010 | 23.2 | 2.7 | 0.02 | 1.4 |
| 2011 | 23.5 | 3.3 | 0.01 | 1.9 |
| 2012 | 24.0 | 1.7 | 0.14 | 0.5 |
| 2013 | 24.0 | 1.7 | 0.22 | 0.5 |
|  | | | | |

SD = standard deviation

p-value from test for between centre variation

Intraclass correlation coefficient (ICC) (%) = 100\*between centre variance/(within centre variance + between centre variance)

**Table 2**

Comparison of mean FEV1% predicted between centres for individual years throughout the study period (unadjusted data)

*3.3 Ranking of Centres*

There was a lack of stability in rankings for the early years of the study but increasing correlation between years as the study progressed, Table 3.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Year | Year | | | | | | |
| 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 |
|  |  |  |  |  |  |  |  |
| 2007 |  | -.122 | .154 | .454 | .417 | .406 | .124 |
| 2008 | *.528* |  | .531 | .271 | .232 | .115 | .096 |
| 2009 | *.426* | *.002* |  | .256 | .387 | .298 | .417 |
| 2010 | *.013* | *.140* | *.143* |  | .**656** | .507 | .420 |
| 2011 | *.024* | *.210* | *.024* | *<0.001* |  | **.834** | **.772** |
| 2012 | *.029* | *.537* | *.087* | *.002* | *<0.001* |  | .**765** |
| 2013 | *.531* | *.613* | *.016* | *.015* | *<0.001* | *<0.001* |  |
|  |  |  |  |  |  |  |  |

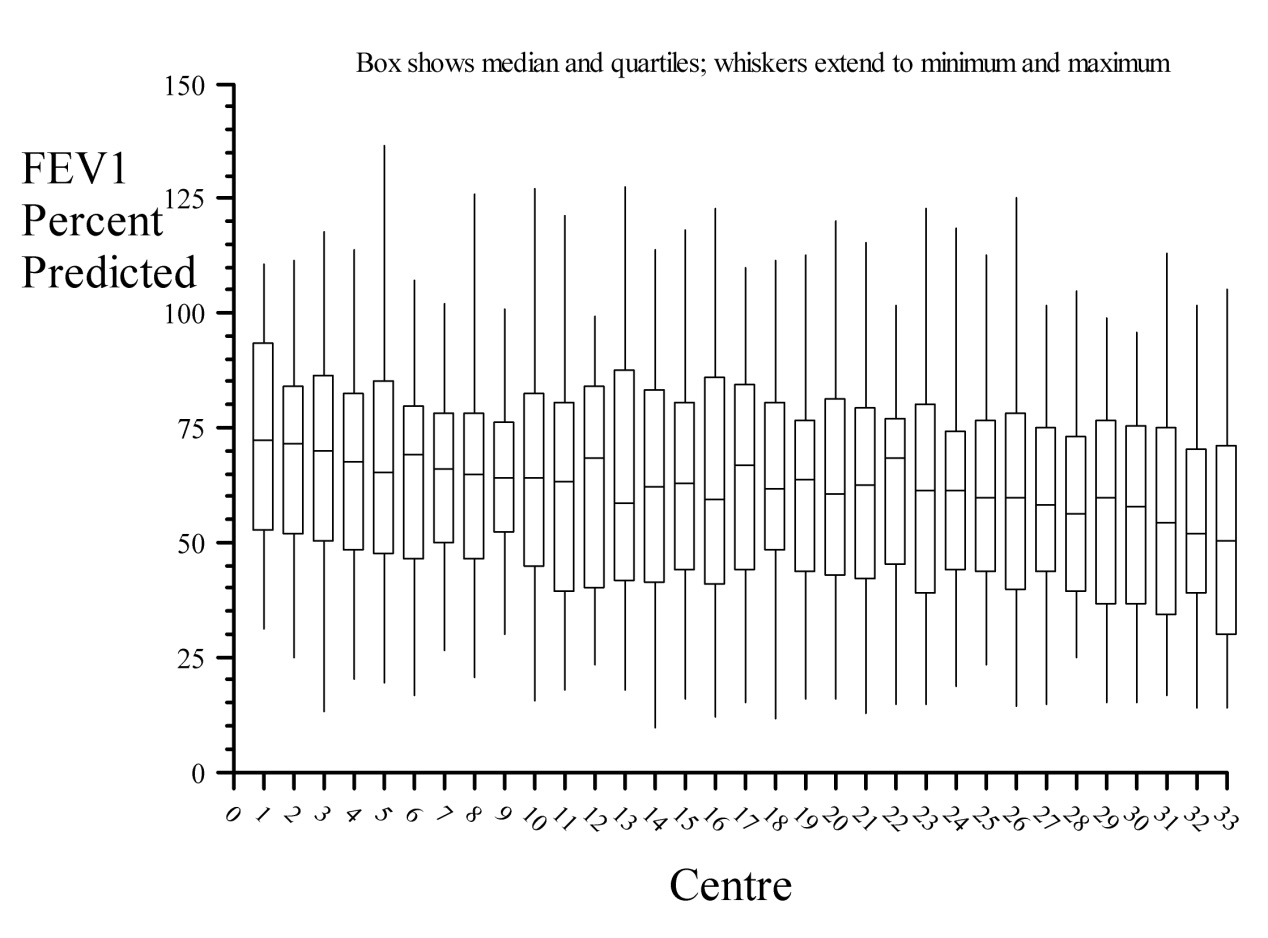
Pearson correlation coefficient shown above the diagonal. Corresponding p-values shown below the diagonal in italics. Values significant at p<0.001 are shown in bold.

**Table 3**

Correlation of centre rankings throughout the study period (based on mean FEV1% predicted)

*3.4 Lung function rate of change*

Presentation of the median FEV1% predicted (as shown in the registry report) leads to a minor re-ordering of the ranking compared to our initial ranking based on mean FEV1% (Figure 1 and Table 4). The use of an alternative lung function measure (change of FEV1% predicted over the study period) leads to marked re-ordering of the ranking of centres (Table 4).



**Figure 1**

2013 data for each of 33 anonymised adult cystic fibrosis centres showing median FEV1% predicted for each centre

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Mean FEV1% predicted 2013 | | Median FEV1% predicted 2013 | | Change in FEV1% 2007-13 (z-score) | | |
| Rank | Mean | Rank | Median | Rank | Mean | 95% CI | |
|  |  |  |  |  |  |  | |
| 1 | 72.30 | 1 | 72.37 | 1 | .389 | .030 to .748 | |
| 2 | 68.54 | 2 | 71.35 | 2 | .330 | .025 to .635 | |
| 3 | 68.34 | 3 | 69.95 | 13 | .053 | -.113 to .219 | |
| 4 | 66.26 | 7 | 67.60 | 10 | .115 | -.001 to .230 | |
| 5 | 65.95 | 10 | 65.07 | 12 | .060 | -.056 to .175 | |
| 6 | 65.33 | 4 | 68.93 | 17 | -.002 | -.249 to .244 | |
| 7 | 65.14 | 9 | 66.09 | 5 | .261 | -.168 to .691 | |
| 8 | 64.77 | 11 | 64.93 | 33 | -.423 | -.673 to -.173 | |
| 9 | 64.63 | 12 | 63.85 | 28 | -.174 | -.430 to .082 | |
| 10 | 63.67 | 13 | 63.85 | 8 | .160 | -.011 to .331 | |
| 11 | 63.20 | 15 | 63.08 | 18 | -.062 | -.235 to .111 | |
| 12 | 63.19 | 5 | 68.29 | 16 | .006 | -.375 to .386 | |
| 13 | 63.11 | 27 | 58.41 | 6 | .167 | -.015 to .349 | |
| 14 | 62.99 | 18 | 61.97 | 15 | .011 | -.147 to .169 | |
| 15 | 62.94 | 16 | 62.81 | 25 | -.117 | -.257 to .022 | |
| 16 | 62.80 | 26 | 59.38 | 20 | -.078 | -.215 to .060 | |
| 17 | 62.75 | 8 | 66.76 | 24 | -.113 | -.291 to .064 | |
| 18 | 62.46 | 19 | 61.69 | 19 | -.065 | -.189 to .060 | |
| 19 | 62.43 | 14 | 63.56 | 9 | .123 | -.029 to .276 | |
| 20 | 62.15 | 22 | 60.69 | 23 | -.099 | -.188 to -.010 | |
| 21 | 61.74 | 17 | 62.32 | 11 | .100 | -.037 to .236 | |
| 22 | 61.51 | 6 | 68.17 | 29 | -.246 | -.502 to .010 | |
| 23 | 60.52 | 21 | 61.19 | 21 | -.094 | -.197 to .009 | |
| 24 | 60.42 | 20 | 61.28 | 14 | .042 | -.142 to .226 | |
| 25 | 60.35 | 25 | 59.61 | 26 | -.154 | -.340 to .033 | |
| 26 | 60.28 | 24 | 59.76 | 7 | .163 | .090 to .236 | |
| 27 | 58.88 | 28 | 58.05 | 32 | -.360 | -.665 to -.054 | |
| 28 | 57.98 | 30 | 56.20 | 4 | .271 | -.044 to .587 | |
| 29 | 57.09 | 23 | 59.77 | 3 | .293 | .009 to .578 | |
| 30 | 55.66 | 29 | 57.89 | 31 | -.268 | -.522 to -.013 | |
| 31 | 55.14 | 31 | 54.31 | 30 | -.254 | -.531 to .024 | |
| 32 | 54.82 | 32 | 51.85 | 22 | -.097 | -.360 to .167 | |
| 33 | 52.02 | 33 | 50.22 | 27 | -.171 | -.495 to .152 | |
|  |  |  |  |  |  |  | |

**Table 4**

Ranking of centres for mean FEV1% predicted , median FEV1% predicted and mean rate of change of FEV1%

*3.5 Controlling for patient characteristics within centres*

The linear mixed model shows that that FEV1% values in 2013 declined with age and were lower in men than women (Table 5). ΔF508 homozygous genotype was associated with a small reduction in FEV1%, but there was a decrease of around 15% in those who were using pancreatic enzymes. Allowing for these patient characteristics reduced the standard deviation of between-centre effects from 1.70 (Table 2, p=0.22) to 0.9 (Table 5, p=0.57) for mean FEV1%.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Regression component | Mean FEV1% predicted (2013 data) | | | | | Change in FEV1% 2007-13 (z-score) | | | | |
| Main effect | Est | SE | CI-Low | CI-High | p-value | Est | SE | CI-Low | CI-High | p-value |
|  |  |  |  |  |  |  |  |  |  |  |
| Intercept | 90.1 | 2.1 | 86.0 | 94.3 | <0.001 | 0.17 | 0.08 | 0.01 | 0.33 | 0.04 |
| Age (years) | -0.38 | 0.04 | -0.45 | -0.30 | <0.001 | 0.004 | 0.002 | 0.001 | 0.007 | 0.02 |
| Sex (1=male, 0=female) | -2.20 | 0.75 | -3.67 | -0.73 | 0.003 | -0.06 | 0.03 | -0.12 | 0.00 | 0.04 |
| ΔF508 homozygous (1=yes, 0=no) | -1.53 | 0.82 | -3.13 | 0.07 | 0.06 | -0.05 | 0.03 | -0.11 | 0.01 | 0.1 |
| Using pancreatic enzymes (1=yes, 0=no) | -14.8 | 1.2 | -17.1 | -12.5 | <0.001 | -0.20 | 0.05 | -0.29 | -0.11 | <0.001 |
| Unclear whether using pancreatic enzymes (1=yes, 0=no) | -8.3 | 1.9 | -12.0 | -4.5 | <0.001 | -0.05 | 0.07 | -0.20 | 0.10 | 0.5 |
|  |  |  |  |  |  |  |  |  |  |  |
| Variance component | Est | SE | SD | ICC | p-value | Est | SE | SD | ICC | p-value |
|  |  |  |  |  |  |  |  |  |  |  |
| Variance between subjects within a centre, τ2 | 541 | 12 | 23.2 | 99.85 | <0.001 | 0.973 | 0.020 | 0.986 | 98.7 | <0.001 |
|  |  |  |  |  |  |  |  |  |  |  |
| Variance between centres, σ2 | 0.80 | 1.40 | 0.90 | 0.15 | 0.57 | 0.013 | 0.006 | 0.114 | 1.3 | 0.04 |

Est = parameter estimate SE = standard error CI = confidence interval

SD = standard deviation ICC = intra-class correlation

**Table 5**

Estimate of components of variance for cross-sectional and longitudinal lung function outcome measures

**4. Discussion**

In the UK, the Cystic Fibrosis Registry publishes an annual report that is publicly available and includes respiratory and nutritional outcome data presented separately for each specialist adult and paediatric centre [9]. Median FEV1 percent predicted is used as the pulmonary outcome measure. Data are presented as a ranked graph and the most recent report also contains funnel plots. Both of these presentations might be interpreted as implying that some centres are ‘better’ than others. We wished to determine whether the published rankings are indicative of standards of care within adult CF units.

We have shown a marginal statistically significant difference in mean FEV1% between centres for some years of the study, which may, at face value, imply a clinically important difference. For example, in 2011, the standard deviation for mean FEV1% was 3.3% suggesting that in this year 95% of centres would have a mean FEV1% contained within a range of 13.2% (four standard deviations). The marginal p-values suggest the results may be influenced by sample size. A power calculation using the entire dataset found that to reliably detect a difference of 5% predicted between centres at the 95% significance level would require a minimum of 273 patients per centre and to detect a 2% difference would require at least 1704 patients in each centre.

Secondly, we wished to determine the stability of rankings. We hypothesized that if the ranking of a centre is related to the standard of care provided, then rankings should not be substantially altered from year to year as the standard of care provided by a centre is unlikely to undergo rapid changes. Additionally, rankings should not be substantially altered if a reasonable alternative outcome measure is employed to rank centres. There is correlation in centre rankings for different years from 2009 onwards which becomes increasingly significant over time. The lack of correlation in earlier years of the study may be related to smaller sample sizes. The number of reports for each year increased during the study period from around 2000 patients in 2007 to around 4000 in 2011. A gradual small increase in patient numbers is expected across centres due to the continually improving survival in cystic fibrosis. However, the increase was not uniform and some centres had a dramatic increase in data over time. Whilst this may represent an influx of new patients in these centres (as can occur when a centre takes over a cohort of patients previously cared for in another hospital) it is more likely that the main influence is improved data entry over the course of the study.

Since the majority of patients die from lung disease some measure of lung function is clearly important to assess care within centres but which measure is the most informative is currently not clear. The UK Registry uses median FEV1% predicted as the outcome measure of choice. There is no recognised ‘gold standard’ for respiratory outcomes in CF but FEV1% is recognised as the most important single factor predictive of survival [14-17]. As such, median FEV1% for a centre would appear a reasonable comparator. We based our initial rankings on mean FEV1% rather than median values to allow us to employ mixed modelling during the remainder of our analyses.

There are some difficulties in the use of cross-sectional data to compare adult centres [18]. Firstly, the median FEV1% of any adult centre will partly be determined by care provided during paediatrics. Adult centres with a good paediatric unit are advantaged by receiving patients with better than average lung function on transition to adult services. Secondly, lung function does not decline in a linear manner until death. Most patients reach a lower plateau of around 20-30% predicted and remain at this level until they die or receive a lung transplant [19]. Centres with better than average care of patients with end stage lung disease would have a lower median FEV1% due to prolonged survival of patients with low lung function. A recent estimate from a single UK centre quoted a median survival of 5.3 years after reaching an FEV1 of less than 30% predicted [20]. Finally, the characteristics of the centre may be important. Larger adult centres and those co-located with a transplant centre may be more likely to attract patients with severe lung disease. In addition, because lung function declines more rapidly in CF than in a healthy population, median FEV1% will be influenced by the age profile of patients within a centre. A newer centre may have a greater proportion of younger patients, who would be expected to have higher FEV1% values, increasing the centre average.

An alternative strategy is to assess rate of decline in lung function for individual patients within a centre. Increased rate of decline in FEV1 is associated with increased mortality [21, 22] and rate of decline of FEV1 has been shown to be a better predictor of death than a single value of FEV1 [23]. Hence we chose the longitudinal measure of rate of decline of FEV1% over the course of the study as our alternative outcome measure.

We have shown that the ranking of centres is critically dependent on the outcome measure chosen. Even the minor change of expressing rankings as means rather than medians leads to some re-ranking of centres. Use of a longitudinal outcome measure leads to marked re-ranking of centres. The most striking example being that the centre ranked 29th for mean FEV1% is ranked third for z-slope score. Since the 2013 data is used to calculate both cross-sectional and longitudinal rankings we would expect there to be some positive correlation between rankings for mean FEV1% in 2013 and a centre’s mean slope over the study period. Despite this, the rankings are substantially altered in many cases suggesting rankings are more influenced by chance than any significant clinical differences between centres.

Although we were able to demonstrate differences in FEV1% between centres for some years of the study, we also found marked centre differences in terms of patient characteristics. There were highly significant differences between centres in terms of genotype and use of pancreatic enzymes and some differences in age. Since all these factors are known to influence disease progression the centre differences will render any direct centre comparison invalid. We found that lung function decline is highly influenced by use of pancreatic enzymes and to a lesser extent by age, sex and genotype. Once data is corrected taking into account these characteristics the difference between centres for pulmonary outcomes is no longer apparent for the 2013 cross-sectional data and barely significant for the longitudinal data. For both outcomes the degree of variation in values derives much more from differing populations within a clinic than from variation between clinics. That is, an individual patient’s lung function is only mildly influenced by the centre they attend. This implies that there is unlikely to be significant differences in the standard of care received by patients at different UK centres.

Our study has the advantage of including a large number of patients with over 4000 CF patients’ data included. This is thought to represent the vast majority of CF patients living in the UK.

Our analyses have a number of limitations. Comparisons between centres include no information about mortality and outcomes could be influenced by sicker subjects dying and leaving a healthier population. An alternative strategy to compare outcomes between centres would be to only include patients within certain age ranges (for example, 20-24 year olds). Previous work by Schluchter et al [24] attempting to phenotype patients according to disease severity has shown that estimated FEV1 at the age of twenty is a good discriminator between patients with mild and severe disease. We found no difference between centres for mean FEV1% when comparisons were made using an age cohort of 20-24 year olds but this may be due to the smaller patient numbers further reducing statistical power.

Additionally, data from different centres may not be equivalent throughout the study. Prior to 2012 there was no stipulation regarding which FEV1% value for the year should be included in the registry for each patient. In some centres the value recorded may have been the patients best in year value whilst other centres may have recorded the value measured on the date of the patient’s annual review. This would have had the effect of improving outcomes for those centres recording best in year values and artificially increasing between centre variance. Since 2012 the registry has stipulated that the value recorded should be the patients best in year value making later centre comparisons fairer. Since our longitudinal data have used data across the whole study period they may have been influenced by this factor.

Comparison between centres will also be affected by centre size. The standard errors for each outcome parameter are generally larger for smaller centres and the centres with the most extreme mean FEV1% values are often the smallest. The UK CF registry has tried to address this issue by introducing the use of funnel plots [10] in the most recent registry report. However, the funnel plots may also be misleading. For example, in the 2015 registry report [9] there is one adult centre that clearly stands out as having markedly higher lung function on the funnel plot for median FEV1% predicted. However, the funnel plot for age shows this centre lies close to 3 standard deviations below the median suggesting that the higher than average lung function is likely to be related at least in part to the age of patients attending the centre.

Finally, there may be other unmeasured factors, such as burden of treatment, for which registry information is not available that influence any apparent difference between centres.

Further comparison of centres in the future is likely to be more informative as the registry now captures a very high proportion of adult CF patients cared for in the UK and since the introduction of best in year FEV1 for the annual dataset in 2012, all centres are now entering equivalent data making comparisons fairer. However, given the marked differences in the characteristics of patients attending centres any comparison that does not adjust for these factors will remain invalid.

**5. Conclusion**

It is possible to demonstrate differences between UK adult CF centres for the mean FEV1% for some years of the study period. However, rankings of centres are critically dependent on the chosen outcome measure and substantially altered when a longitudinal measure is employed. In addition, there are marked differences between centres for the characteristics of their patients and variation between centres becomes negligible once these characteristics are taken into account. This implies that any apparent difference in respiratory outcomes is unlikely to be related to differing standards of care between centres.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Centre ranking for 2013 | Mean number of observations | Year | | | | | | | Total |
| 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 |
|  |  |  |  |  |  |  |  |  |  |
| 1 | 4.0 | 2 | 6 | 17 | 14 | 22 | 23 | 25 | 109 |
| 2 | 5.1 | 24 | 28 | 32 | 28 | 32 | 36 | 35 | 215 |
| 3 | 5.1 | 75 | 95 | 105 | 111 | 124 | 125 | 122 | 757 |
| 4 | 5.1 | 109 | 132 | 163 | 173 | 185 | 173 | 173 | 1108 |
| 5 | 5.6 | 169 | 198 | 217 | 232 | 239 | 224 | 227 | 1506 |
| 6 | 4.3 | 18 | 21 | 21 | 28 | 38 | 21 | 33 | 180 |
| 7 | 4.0 | 0 | 0 | 2 | 10 | 12 | 12 | 12 | 48 |
| 8 | 5.1 | 35 | 38 | 47 | 57 | 64 | 70 | 69 | 380 |
| 9 | 3.9 | 0 | 0 | 11 | 13 | 14 | 14 | 15 | 67 |
| 10 | 3.6 | 0 | 1 | 93 | 30 | 97 | 126 | 122 | 469 |
| 11 | 5.0 | 84 | 102 | 108 | 103 | 111 | 115 | 110 | 733 |
| 12 | 4.9 | 5 | 10 | 14 | 18 | 22 | 22 | 22 | 113 |
| 13 | 4.3 | 28 | 45 | 27 | 42 | 59 | 63 | 62 | 326 |
| 14 | 4.3 | 51 | 51 | 68 | 103 | 140 | 143 | 142 | 698 |
| 15 | 5.1 | 86 | 211 | 232 | 214 | 262 | 267 | 261 | 1533 |
| 16 | 5.4 | 129 | 172 | 185 | 197 | 207 | 225 | 208 | 1323 |
| 17 | 4.6 | 35 | 68 | 74 | 61 | 91 | 92 | 99 | 520 |
| 18 | 4.9 | 125 | 125 | 177 | 177 | 225 | 232 | 203 | 1264 |
| 19 | 4.7 | 45 | 55 | 94 | 106 | 123 | 128 | 127 | 678 |
| 20 | 5.1 | 147 | 213 | 217 | 290 | 300 | 320 | 321 | 1808 |
| 21 | 3.6 | 0 | 14 | 2 | 164 | 179 | 195 | 186 | 740 |
| 22 | 3.6 | 1 | 2 | 1 | 26 | 38 | 39 | 39 | 146 |
| 23 | 5.3 | 35 | 309 | 321 | 340 | 347 | 353 | 341 | 2046 |
| 24 | 5.4 | 44 | 60 | 71 | 82 | 82 | 85 | 74 | 498 |
| 25 | 5.2 | 83 | 87 | 90 | 92 | 101 | 99 | 92 | 644 |
| 26 | 5.4 | 402 | 393 | 509 | 525 | 552 | 582 | 556 | 3519 |
| 27 | 4.8 | 12 | 20 | 23 | 23 | 27 | 28 | 26 | 159 |
| 28 | 4.8 | 12 | 15 | 17 | 18 | 21 | 20 | 18 | 121 |
| 29 | 4.7 | 13 | 17 | 31 | 39 | 36 | 43 | 42 | 221 |
| 30 | 5.1 | 33 | 43 | 38 | 46 | 52 | 53 | 50 | 315 |
| 31 | 4.9 | 40 | 39 | 37 | 38 | 45 | 38 | 47 | 284 |
| 32 | 5.0 | 30 | 34 | 32 | 36 | 38 | 44 | 42 | 256 |
| 33 | 3.2 | 0 | 0 | 23 | 2 | 31 | 28 | 26 | 110 |
|  |  |  |  |  |  |  |  |  |  |
| Unranked | 3.8 | 84 | 118 | 102 | 92 | 116 | 27 | 0 | 539 |
|  |  |  |  |  |  |  |  |  |  |
| Annual total |  | 1956 | 2722 | 3201 | 3530 | 4032 | 4065 | 3927 | 23433 |
|  |  |  |  |  |  |  |  |  |  |

**Supplementary Table**

Number of reports by year (clinics ranked by mean FEV1% predicted for 2013)

**ACKNOWLEDGEMENTS**

We would like to acknowledge the support of the UK CF Trust that made this analysis possible. We would also like to thank all the CF patients who consent to their data being part of the UK CF Registry.

**COMPETING INTERESTS**

None declared.

**FUNDING**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**CONTRIBUTORS**

JAN and CO contributed to the conception and design of the study. CO performed the statistical analysis and JAN and CO contributed to the interpretation of the data. JAN drafted the manuscript and CO revised the article critically. Both authors contributed to the final version of the article and approval of the version to be published.

**ETHICS APPROVAL**

NHS research ethics approval was granted for UK CF Registry and each patient or legally authorised representative provided written informed consent for data collection and research. Under the terms of the NHS ethics approval, the UK CF trust steering committee approved this study.

**References**

[1] [Salvatore D](http://www.ncbi.nlm.nih.gov/pubmed/?term=Salvatore%20D%5BAuthor%5D&cauthor=true&cauthor_uid=22884375), [Buzzetti R](http://www.ncbi.nlm.nih.gov/pubmed/?term=Buzzetti%20R%5BAuthor%5D&cauthor=true&cauthor_uid=22884375), [Baldo E](http://www.ncbi.nlm.nih.gov/pubmed/?term=Baldo%20E%5BAuthor%5D&cauthor=true&cauthor_uid=22884375), Furnari ML, Lucidi V, Manunza D *et al.* An overview of international literature from cystic fibrosis registries. Part 4: update 2011. [*J Cyst Fibros*](http://www.ncbi.nlm.nih.gov/pubmed/22884375) 2012;11(6):480-93.

[2] Stern M, Niemann N, Wiedemann B, Wenzlaff P on behalf of the German CFQA Group. Benchmarking improves quality in cystic fibrosis care: a pilot project involving 12 centres. [*Int J Qual Health Care*](http://www.ncbi.nlm.nih.gov/pubmed/21504960)2011;23(3):349-56.

[3] Boyle MP, Sabadosa KA, Quinton HB, Marshall BC, Schechter MS. Key findings of the US Cystic Fibrosis Foundation's clinical practice benchmarking project. [*BMJ Qual Saf*](http://www.ncbi.nlm.nih.gov/pubmed/24608546) 2014;23:i15-i22.

[4] [Martin B](http://www.ncbi.nlm.nih.gov/pubmed/?term=Martin%20B%5BAuthor%5D&cauthor=true&cauthor_uid=22250024), [Schechter MS](http://www.ncbi.nlm.nih.gov/pubmed/?term=Schechter%20MS%5BAuthor%5D&cauthor=true&cauthor_uid=22250024), [Jaffe A](http://www.ncbi.nlm.nih.gov/pubmed/?term=Jaffe%20A%5BAuthor%5D&cauthor=true&cauthor_uid=22250024), Cooper P, Bell SC, Ranganathan S. Comparison of the US and Australian cystic fibrosis registries: the impact of newborn screening. *Pediatrics* 2012;129(2):e348-55.

[5] [Goss](http://thorax.bmj.com/search?author1=Christopher+H+Goss&sortspec=date&submit=Submit) CH, [MacNeill](http://thorax.bmj.com/search?author1=Stephanie+J+MacNeill&sortspec=date&submit=Submit) SJ, [Quinton](http://thorax.bmj.com/search?author1=Hebe+B+Quinton&sortspec=date&submit=Submit) HB, Marshall BC, Elbert A, Knapp EA *et al*. Children and young adults with CF in the USA have better lung function compared with the UK. *Thorax* 2015;70:229-36.

[6] Corey M, McLaughlin FJ, Williams M, Levison H. A comparison of survival, growth, and pulmonary function in patients with cystic fibrosis in Boston and Toronto. *J Clinical Epidemiology* 1988;41:583-91.

[7] Johnson C, Butler SM, Konstan MW, Morgan W, Wohl ME. Factors influencing outcomes in cystic fibrosis: a center-based analysis. *Chest* 2003;123(1):20-7.

[8] [Padman](http://pediatrics.aappublications.org/search?author1=Raj+Padman&sortspec=date&submit=Submit) R, [McColley](http://pediatrics.aappublications.org/search?author1=Susanna+A.+McColley&sortspec=date&submit=Submit) SA,  [Miller](http://pediatrics.aappublications.org/search?author1=Dave+P.+Miller&sortspec=date&submit=Submit) DP,  [Konstan MW, Morgan WJ, Schechter MS *et al*](http://pediatrics.aappublications.org/search?author1=Michael+W.+Konstan&sortspec=date&submit=Submit) for the Investigators and Coordinators of the Epidemiologic Study of Cystic Fibrosis. Infant Care Patterns at Epidemiologic Study of Cystic Fibrosis Sites That Achieve Superior Childhood Lung Function. *Pediatrics* 2007;119;e531-7.

[9] Cystic fibrosis strength in numbers. UK Cystic Fibrosis Registry 2015 Annual Data Report. Published August 2016.

[10] Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Statistics in Medicine* 2005; 24:1185-1202.

[11] [dataset] Cystic Fibrosis Trust. UK CF Registry. 2013.

[12] 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* 2012;40(6):1324-43.

[13] https://en.wikipedia.org/wiki/Normal\_score.

[14] Kerem E, Reisman J, Corey M, Canny GJ, Levison H. Prediction of mortality in patients with cystic fibrosis. *N Engl J Med* 1992;326(18):1187-91.

[15] Liou TG, Adler FR, Fitzsimmons SC, Cahill BC, Hibbs JR, Marshall BC. Predictive 5-year survivorship model of cystic fibrosis. *Am J Epidemiology* 2001;153(4):345-52.

[16] Schluchter MD, Konstan MW, Davis PB. Jointly modelling the relationship between survival and pulmonary function in cystic fibrosis patients. *Statist Med* 2002;21:1271-87.

[17] Taussig LM, Kattwinkel J, Friedewald WT, Di Sant’Agnese PA. A new prognostic score and clinical evaluation system for cystic fibrosis. *J Pediatrics* 1973;82(3):380-90.

[18] Schechter MS. Benchmarking to improve the quality of cystic fibrosis care. *Curr Opin Pulm Med* 2012;18(6):596-601.

[19] Konstan MW, Wagener JS, VanDevanter DR, Pasta DJ, Yegin A, Rasouliyan L et al. Risk factors for rate of decline in FEV1 in adults with cystic fibrosis. J Cyst Fibros 2012;11(5):405-11.

[20] George PM, Banya W, Pareek N, Bilton D, Cullinan P, Hodson ME *et al*. Improved survival at low lung function in cystic fibrosis: cohort study from 1990 to 2007. *BMJ* 2011;342:d1008.

[21] Schluchter MD, Konstan MW, Davis PB. Jointly modelling the relationship between survival and pulmonary function in cystic fibrosis patients. *Statist Med* 2002;21:1271-87.

[22] Corey M, Edwards L, Levison H, Knowles M. Longitudinal analysis of pulmonary function decline in patients with cystic fibrosis. *J Pediatrics* 1997;131:809-14.

[23] Milla CE, Warwick WJ. Risk of death in cystic fibrosis patients with severely compromised lung function. *Chest* 1998;113:1230-4.

[24] Schlucter MD, Konstan MW, Drumm ML, Yankaskas JR, Knowles MR. Classifying severity of cystic fibrosis lung disease using longitudinal pulmonary function data. *Am J Respir Crit Care Med* 2006;174:780-6.