**Appendix A: Further explanation and details regarding the age-adjustment for % FEV1 using a Generalised Linear Model (GLM) approach**

The first step was to explore the age data year by year and divide the cohort into 10 age groups with roughly similar number of people in each group. The following age ranges were chosen empirically following exploration of the data:

|  |  |  |
| --- | --- | --- |
| Number of people (%) | **2013 (n = 4269)** | **2014 (n = 4644)** |
| **Age range:**16 – 19 years20 – 21 years22 – 23 years24 – 25 years26 – 27 years28 – 30 years31 – 33 years34 – 37 years38 – 44 years≥45 years | 472 (11.1%)437 (10.2%)444 (10.4%)412 (9.7%)361 (8.5%)465 (10.9%)431 (10.1%)365 (8.6%)422 (9.9%)460 (10.8%) | 504 (10.9%)459 (9.9%)447 (9.6%)433 (9.3%)403 (8.7%)509 (11.0%)453 (9.8%)429 (9.2%)485 (10.4%)522 (11.2%) |

The second step was to calculate the predicted % FEV1 for each age decile using a linear model, with the age deciles as a fixed factor and main effect for order. Although a linear model was used to fit age decile, the actual age adjustment for %FEV1 was non-linear because actual age was transformed into 10 age deciles with varying age ranges in each decile. Data transformation was used to account for the non-linear relationship between age and %FEV1 among adults with CF in this Generalised Linear Model approach. In other words, fitting age decile as a categorical variable in a linear model accounts for the non-linear relationship between %FEV1 and actual age in years. The relationship between %FEV1 and actual age in years is displayed in Figure 1E in Appendix E.

|  |  |  |
| --- | --- | --- |
| Predicted % FEV1 using GLM | **2013****(n = 4269)** | **2014****(n = 4644)** |
| **Age range:**16 – 19 years20 – 21 years22 – 23 years24 – 25 years26 – 27 years28 – 30 years31 – 33 years34 – 37 years38 – 44 years≥45 years | 73.11%69.30%67.15%64.17%63.65%63.95%59.95%61.82%62.86%62.16% | 75.42%72.01%66.87%67.14%62.64%64.91%61.30%62.46%63.06%61.54% |

The third step was to calculate the age-adjusted FEV1 for every study subject, by calculating the percentage of actual %FEV1 over the predicted %FEV1. A % age-adjusted FEV1 of 100% meant the person’s %FEV1 was as expected for his / her age. A % age-adjusted FEV1 >100% meant the person’s %FEV1 was better than expected for his / her age. A % age-adjusted FEV1 <100% meant the person’s %FEV1 was worse than expected for his / her age.

There were significant differences in the age of adults with CF between different centres. Therefore, this age adjustment is crucial to ensure that the ranking of the CF centres are robust to the confounding effects of case-mix. More detailed analyses methods are discussed in Appendix C and robustness of the ranking process is discussed in Appendix E.

**Appendix B: Methods and results for regression modelling of IV use and preventive therapies data**

The 1995-1996 ESCF analysis used stratification method to control for confounding, but regression modelling was used in the 2003-2005 paediatric ESCF analysis [1]. In the 2003-3005 ESCF analysis, antibiotic treatment was adjusted for gender, race / ethnicity, age, Medicaid status, baseline disease severity (FEV1 in study subjects over six years) and *P. aeruginosa* status [1]. Within the UK, universal healthcare is available via the NHS, hence private insurance status is extremely unlikely to affect the use of IV antibiotics. There are only a small number of non-Caucasian people with CF; hence ethnicity is unlikely to influence the results. IV use were therefore adjusted for gender, age, pancreatic status (as a surrogate for people with ‘typical’ vs ‘mild phenotype’ CF), %FEV1, and *P. aeruginosa* status in this regression modelling analysis.

The adjustment of IV days (as a continuous variable) for the five case-mix factors mentioned above was done in our analysis using a similar Generalised Linear Model (GLM) approach for age-adjustment of %FEV1 (see Appendix A). This involved calculating the predicted IV days for each study subject using a linear model, with age deciles (same age deciles in Appendix A), pancreatic status (on pancreatic enzyme replacement therapy i.e. pancreatic insufficient vs not on pancreatic enzyme replacement therapy i.e. pancreatic sufficient), gender (male vs female), %FEV1 (<40% vs 40-69.9% vs ≥70%) and *P. aeruginosa* status (chronic vs intermittent vs no *P. aeruginosa*) as fixed factors and main effect for order. Fitting age deciles and %FEV1 as categorical variables allow for the fact that the relationships between IV days with age, and IV days with %FEV1 were non-linear. The next step was to calculate the case-mix adjusted IV days for every individual, by subtracting the predicted IV days from the actual IV days. A case-mix adjusted IV days of 0 meant the person’s IV days was as expected for his / her age, gender, pancreatic status, %FEV1 and *P. aeruginosa* status. A case-mix adjusted IV days >0 meant the person had higher IV days than expected for his / her age, gender, pancreatic status, %FEV1 and *P. aeruginosa* status. A case-mix adjusted IV days <0 meant the person had lower IV days than expected for his / her age, gender, pancreatic status, %FEV1 and *P. aeruginosa* status. Adjustment for number of IV courses was done using the same method as adjustment for IV days.

Kruskal-Wallis H test was used to compare IV use between all three groups of specialist adult CF centres (‘upper quarter’, ‘middle half’, ‘lower quarter’). To determine whether results using stratification method (see main text Tables 3 & 4) were similar with to the results using regression modelling, Mann-Whitney U test was used to compare case-mix adjusted IV use between ‘upper quarter’ vs ‘lower quarter’, ‘upper quarter’ vs ‘middle half’ and ‘middle half’ vs ‘lower quarter’, with Bonferroni correction applied for multiple comparisons. Non-parametric tests were used in these comparisons because case-mix adjusted IV days and case-mix adjusted number of IV courses were both remained skewed (albeit less skewed compared to unadjusted IV days or unadjusted number of IV courses) despite adjustment using the Generalised Linear Model approach.

Table 1B: Annual case-mix adjusted number of IV courses and annual case-mix adjusted IV antibiotic days for the three groups of specialist CF centres for 2013 and 2014

|  |  |  |
| --- | --- | --- |
|  | 2013 | 2014 |
| Upper quarter(n = 981) | Middle half(n = 1867) | Lower quarter(n = 1244) | Upper quarter(n = 871) | Middle half(n = 2167) | Lower quarter(n = 1411) |
| Annual case-mix adjusted IV courses,median (IQR) † | –0.29 (–0.87 to 0.62) | –0.25 (–0.99 to 0.92) | –0.40 (–1.21 to –0.39) | –0.24 (–0.93 to 0.61) | –0.25 (–0.96 to 0.84) | –0.45 (–1.18 to 0.38) |
| Annual case-mix adjusted IV days,median (IQR) ‡ | –4.8 (–13.2 to 8.6) | –4.3 (–15.5 to 12.1) | –5.9 (–19.2 to 6.3) | –3.9 (–13.3 to 8.3) | –4.3 (–14.9 to 10.0) | –6.5 (–18.8 to 3.9) |

† Median adjusted IV course of –0.29 for upper quarter in 2013 meant the group as a whole has a median of 0.29 fewer IV courses in that year compared to what was expected for the case-mix of that group. That indicated higher number of IV courses used compared to a median adjusted IV course of –0.40 for lower quarter.

‡ Median adjusted IV days of –4.8 for upper quarter in 2013 meant the group as a whole has a median of 4.8 fewer days of IV in that year compared to what was expected for the case-mix of that group. That indicated higher IV days compared to a median adjusted IV days of –5.9 for lower quarter.

For 2013, for case-mix adjusted number of IV courses

P-value§ for comparison of all three groups < 0.001

**P-valueΩ for upper vs lower quarter** < 0.001 (after Bonferroni correction, p-value **< 0.001**)

**P-valueΩ for middle half vs lower quarter** < 0.001 (after Bonferroni correction, p-value **< 0.001**)

P-valueΩ for upper quarter vs middle half = 0.842 (after Bonferroni correction, p-value = 1)

For 2013, for case-mix adjusted IV days

P-value§ for comparison of all three groups < 0.001

**P-valueΩ for upper vs lower quarter** < 0.001 (after Bonferroni correction, p-value **< 0.001**)

**P-valueΩ for middle half vs lower quarter** < 0.001 (after Bonferroni correction, p-value **< 0.001**)

P-valueΩ for upper quarter vs middle half = 0.546 (after Bonferroni correction, p-value = 1)

For 2014, for case-mix adjusted number of IV courses

P-value§ for comparison of all three groups < 0.001

**P-valueΩ for upper vs lower quarter** < 0.001 (after Bonferroni correction, p-value **< 0.001**)

**P-valueΩ for middle half vs lower quarter** < 0.001 (after Bonferroni correction, p-value **< 0.001**)

P-valueΩ for upper quarter vs middle half = 0.934 (after Bonferroni correction, p-value = 1)

For 2014, for case-mix adjusted IV days

P-value§ for comparison of all three groups < 0.001

**P-valueΩ for upper vs lower quarter** < 0.001 (after Bonferroni correction, p-value **< 0.001**)

**P-valueΩ for middle half vs lower quarter** < 0.001 (after Bonferroni correction, p-value **< 0.001**)

P-valueΩ for upper quarter vs middle half = 0.734 (after Bonferroni correction, p-value = 1)

§ Kruskal-Wallis H test was used to compare all three groups Ω Mann-Whitney U test was used to compare two groups

Table 1B in the previous page summarises the case-mix adjusted IV use for the three groups of CF centres, and the results of all the comparisons. There were statistically significant differences in both case-mix adjusted IV days and case-mix adjusted number of IV courses for those three groups of specialist adult CF centres (‘upper quarter’, ‘middle half’, ‘lower quarter’). However, ‘upper quarter’ had similar IV use when compared with ‘middle half’. The between-group differences were primarily driven by lower IV antibiotics use in the ‘lower quarter’. Results for case-mix adjusted IV days were consistent with results for case-mix adjusted number of IV courses. Results in 2013 were also entirely consistent with results in 2014.

In the main analysis (see section 2.2 of the main text), IV use was also analysed using stratification method as a binary variable (i.e. the proportion of people prescribed at least one IV course per year). To mirror this analysis, we also used regression modelling to analyse the proportion of people prescribed at least one IV course per year. For this analysis of IV use as a binary variable, (prescribed at least one course of IV antibiotics per year vs not prescribed any), we used a similar method to the 2003-2005 ESCF analyses [1]. As discussed earlier, we adjusted IV use for gender, age, pancreatic status (as a surrogate for people with ‘typical’ vs ‘mild phenotype’ CF), %FEV1, and *P. aeruginosa* status in a regression modelling analysis. The same categorical variables were therefore included in a binary logistic model to do a pairwise comparison for the odds of being prescribed at least one course of IV antibiotics per year at the upper quarter centres vs middle half centres vs lower quarter centres.

Table 2B: Adjusted odds ratio for being prescribed at least one course of IV antibiotics per year according to the three groups of specialist CF centres for 2013 and 2014

|  |  |  |
| --- | --- | --- |
|  | 2013 † | 2014 ‡ |
| Adjusted\* OR (95% CI) | P-value | Adjusted\* OR (95% CI) | P-value |
| ‘Upper quarter’ vs ‘lower quarter’ (ref) | **1.48 (1.21 – 1.80)** | **< 0.001** | **1.37 (1.12 – 1.67)** | **0.002** |
| ‘Middle half’ vs ‘lower quarter’ (ref) | **1.53 (1.29 – 1.81)** | **< 0.001** | **1.29 (1.10 – 1.51)** | **0.002** |
| ‘Upper quarter’ vs ‘Middle half’ (ref) | 0.97 (0.81 – 1.16) | 0.721 | 1.06 (0.88 – 1.28) | 0.518 |

† For this logistic model: pseudo-R2 = 0.317 (Nagelkerke); model *χ*2(17) = 1090.9, p <0.001

‡ For this logistic model: pseudo-R2 = 0.324 (Nagelkerke); model *χ*2(17) = 1223.5, p <0.001

\*Adjusted for gender, age decile, pancreatic status, %FEV1 categories and *P. aeruginosa* status

These results are consistent with the results in Table 1B. ‘Upper quarter’ had similar IV use in comparison to ‘middle half’, but ‘lower quarter’ had lower IV antibiotics use. Therefore, the results for IV use were very similar for both regression modelling and stratification method (see Tables 3-5 in the main text). This similarity demonstrates the robustness of the IV use results in this study.

In this study, prescription data for preventive therapies (inhaled antibiotics, inhaled mucolytics and long-term oral macrolide) were also analysed. These were binary variables (prescribed vs not prescribed) – see Appendix G for further details regarding the preventive therapies data. The preventive therapies data were not accompanied by any adherence data; hence it is difficult to understand how much of the prescribed therapy was actually used. On the other hand, if the registry data suggested that a preventive therapy was not prescribed, it is likely the treatment had not been used. We could be relatively confident that adults not prescribed any preventive therapies were not using any of those therapies, hence they were different from those who were prescribed at least one type of preventive therapies. This allows some sort of interpretation using the registry data for preventive therapy. In the main analysis of the preventive therapies prescription data, study subjects were therefore divided into two groups – those prescribed at least one form of preventive therapies (inhaled antibiotics / inhaled mucolytics / long-term oral macrolide) vs those who were not prescribed any preventive therapies

In the main analysis of the preventive therapies prescription data, stratification method was used (results in Table 6 of the main text). To mirror the analyses we have performed for IV use, we also used regression modelling to analyse the preventive therapies prescription data. The 2003-2005 ESCF analyses used a binary logistic regression model to analyse a binary outcome (antibiotics prescribed vs not prescribed) [1]. Since our preventive therapy “outcome” was also binary (prescribed at least one form of preventive therapy vs not prescribed any), we used a similar method to the 2003-2005 ESCF analyses. As discussed earlier, we adjusted IV use for gender, age, pancreatic status (as a surrogate for people with ‘typical’ vs ‘mild phenotype’ CF), %FEV1, and *P. aeruginosa* status in a regression modelling analysis. The same categorical variables were therefore included in a binary logistic model to do a pairwise comparison for the odds of being prescribed at least one form of preventive therapy at the upper quarter centres vs middle half centres vs lower quarter centres.

The results of this regression modelling analyses are summarised in Table 3B (see next page). For both 2013 and 2014, the ‘middle half’ centres were least likely to prescribe at least one form of preventive therapies, although this was not statistically significant on 2013. In 2013, ‘upper quarter’ centres were more likely to prescribed at least one form of preventive therapies compared to ‘lower quarter’ and ‘middle half’. However, this was inconsistent with the results in 2014, whereby ‘lower quarter’ centres were just as likely to prescribe at least one form of preventive therapy in comparison to ‘upper quarter’ centres. ‘Middle half’ centres were the least likely to prescribe at least one form of preventive therapies in 2014. This does not explain the between-group differences in FEV1, because lower prescription of preventive therapies should not improve FEV1 of ‘middle half’ (compared to the ‘lower quarter’ in 2014) whilst at the same time reduce FEV1 of the ‘middle half’ (compared to the ‘higher quarter’ in 2014).

Table 3B: Adjusted odds ratio for being prescribed at least one type of pulmonary preventive therapies according to the three groups of specialist CF centres for 2013 and 2014

|  |  |  |
| --- | --- | --- |
|  | 2013 † | 2014 ‡ |
| Adjusted\* OR (95% CI) | P-value | Adjusted\* OR (95% CI) | P-value |
| ‘Upper quarter’ vs ‘lower quarter’ (ref) | **1.45 (1.07 – 1.95)** | **0.016** | 0.94 (0.67 – 1.32) | 0.730 |
| ‘Middle half’ vs ‘lower quarter’ (ref) | 0.94 (0.73 – 1.21) | 0.632 | **0.49 (0.37 – 0.65)** | **< 0.001** |
| ‘Upper quarter’ vs ‘Middle half’ (ref) | **1.54 (1.16 – 2.03)** | **0.003** | **1.92 (1.43 – 2.59)** | **< 0.001** |

† For this logistic model: pseudo-R2 = 0.362 (Nagelkerke); model *χ*2(17) = 866.3, p <0.001

‡ For this logistic model: pseudo-R2 = 0.412 (Nagelkerke); model *χ*2(17) = 1017.0, p <0.001

\*Adjusted for gender, age decile, pancreatic status, %FEV1 categories and *P. aeruginosa* status

Although preventive therapy prescription results from regression modelling are similar to results using stratification methods (see Table 6 of the main text), there are inconsistencies in the results from year-to-year, and in relation to the FEV1 outcomes between groups. This contrasted with the strong and consistent relationship that centres with lower FEV1 have lower IV use. Therefore, it is much more likely for the between-group differences in IV use to explain the differences in FEV1 outcomes, rather than between-group differences in prescription of preventive therapies. Further analyses of the preventive therapies data are available in Appendix G.

REFERENCES:

1. Schechter MS, Regelmann WE, Sawicki GS, et al. Antibiotic treatment of signs and symptoms of pulmonary exacerbations: a comparison by care site. *Pediatr Pulmonol* 2015;50:431-40.

**Appendix C: Further explanation of the data analysis method**

In this analysis of the UK CF registry data, it is important to consider whether the annual review FEV1 systematically under reports the actual FEV1 and whether this bias can be predicted by the size of the CF centre attended by an adult with CF. We have demonstrated that the magnitude of discrepancy between annual review and best annual FEV1 readings within the UK CF registry is a surrogate for the proportion of clinically stable annual reviews – smaller discrepancy indicates a higher proportion of annual review performed during periods of stability and vice versa [1]. Our data suggest that a paired mean difference between annual review and best annual FEV1 of –3.1% (95% CI –3.8% to –2.4%) if all annual reviews were only performed when clinically stable [1].

The overall paired mean difference between annual review and best annual FEV1 within the UK CF registry was –5.6% (95% CI –5.9% to –5.4%), suggesting that not all annual review FEV1 were performed during periods of stability.1 Crucially, the discrepancy between annual review and best annual FEV1 readings also vary from centre-to-centre, with larger discrepancies among smaller and larger centres:

Figure 1C: Graph showing the discrepancy in % FEV1 according to the size of CF centres in 2014

The Local Polynomial Regression (LOESS) curve is a non-parametric method for fitting smooth curves to empirical data, to depict relationships between variables [2]. Figure 1C shows an inverted U-shape relationship between discrepancy in %FEV1. A quadratic regression coefficient is statistically significant, compared with a linear one; p = 0.048 for the quadratic coefficient (*y = b0 + b1x +b2x*2 where y = discrepancy between annual review vs best annual % FEV1, x = number of adults in each centre).

However, there were no clear between-group differences in the discrepancy of %FEV1 when CF centres are aggregated into three groups based on ranking using median % age-adjusted annual review FEV1 (‘upper quarter’, ‘middle half’ and ‘lower quarter’). This is because each group contains CF centres with different sizes; hence discrepancies at the extreme of centre sizes are cancelled out by the discrepancies in the other direction at the middle:

Table 1C: Table showing the discrepancy in %FEV1 for the three groups of CF centres

|  |  |  |
| --- | --- | --- |
|  | 2013 | 2014 |
| Upper quarter\*(n = 185)6 centres | Middle half\*(n = 389)8 centres | Lower quarter\*(n = 350)5 centres | Upper quarter\*(n = 614)6 centres | Middle half\*(n = 1772)13 centres | Lower quarter\*(n = 609)7 centres |
| Annual review % FEV1 vs best annual % FEV1, mean (95% CI) | –6.4 (–7.6 to –5.3) | –7.1 (–7.8 to –6.3) | –6.9 (–7.6 to –6.1) | –6.0 (–6.6 to –5.4) | –5.6 (–5.9 to –5.2) | –5.3 (–5.9 to –4.8) |
| ANOVA p-value | 0.643 | 0.246 |

\* Not all centres supplied best annual FEV1 data.

\*\* Best annual FEV1 data were only available for 924/4269 (21.6%) adults in 2013 and 2995/4644 (64.5%) adults in 2014. Therefore, best annual FEV1 data could not be used to rank CF centres.

These results suggest that a centre-by-centre comparison is vulnerable to systematic bias in annual review FEV1, and it is uncertain if one centre genuinely has better outcomes when directly compared to another centre. This is in part due to problems of trying to detect differences in relatively small samples (see paragraph below on sample size). However, aggregating centres into three large groups has helped to distribute the bias in annual review FEV1 equally between groups. One could be sure that one group of centres have different outcomes to another group of centres (see the analyses using annual review %FEV1 and also best annual %FEV1 stratified according to age in Table 2 of the main text and Appendix E). This is one clear advantage of using the “ESCF method” of aggregating several centres into larger groups for analysis, instead of using a multi-level model to compare IV use between centres.

Another advantage of aggregating CF centres is the sample size needed to detect a difference between centres. Previous power calculation using the UK CF registry data suggests that at least 273 adults per centre are needed to reliably detect a 5% difference in FEV1 at the 95% significance level [3], yet only 5/27 (18.5%) of the centres in 2013 and 6/28 (21.4%) of the centres in 2014 have that number of adults with CF. By aggregating CF centres into three larger groups, each group has ≥944 adults which allows more realistic detection of FEV1 differences between groups.

REFERENCES:

1. Hoo ZH, Curley R, Campbell MJ, Walters SJ, Wildman MJ. Understanding FEV1 for the purpose of cystic fibrosis registry comparisons: are annual review FEV1 in the UK only collected when subjects are well? Submitted to *ERJ Open Res*
2. Cleveland WS, Devlin SJ. Locally weighted regression: an approach to regression analysis by local fitting. *J Am Stat Assoc* 1988; 83: 596–610.
3. Nightingale JA, Osmond C. Does current reporting of lung function by the UK cystic fibrosis registry allow a fair comparison of adult centres? *J Cyst Fibros* 2017 [Epub ahead of print]

**Appendix D: Summary of the number of adults included in the analyses and details regarding missing data**

2014

2013

Number of adults receiving care at specialist adult CF centres

n = 4825

Number of adults receiving care at specialist adult CF centres

n = 5255

Missing data:

Age

Gender

BMI

Annual review % FEV1

IV use (courses & IV days)

Inhaled antibiotics

Inhaled mucolytics

Long-term macrolide

Number of adults included in the analyses

n = 4644

0

0

108 (2.3%)

195 (4.2%)

0

0

0

0

Number of adults included in the analyses

n = 4269

0

0

122 (2.9%)

173 (4.1%)

5 (0.1%)

0

0

0

n = 330

n = 297

n = 281

n = 259

Number of adults on ivacaftor

Number of adults with history of lung transplantation

**Appendix E: Robustness of the ranking process using median % age-adjusted FEV1 in identifying CF centres with better outcomes**

Whilst centres in the ‘lower quarter’ have adults that are slightly older, it should be noted that specialist adult centres in the UK were established at different times and therefore the opportunity to “accumulate” adults with milder phenotypes and better survival differs from centre to centre. Some of the UK adult CF centres are established 2-3 decades before other newer adult CF centres. Due to this discrepancy, age of adults in a centre is not necessarily a marker of differences in survival and it is also not a sensitive marker for the quality of care provided. The stepwise increases in %FEV1 from ‘lower quarter’ to ‘middle half’ to ‘upper quarter’ as shown in Table 2 of the main text were highly significant, indicating that centres in the ‘upper quarter’ have the best outcomes while centres in the ‘lower quarter’ have the worst outcomes.

The differences in FEV1 between the three groups of CF centres among those aged 16–19 years should not be interpreted as centres in the ‘lower quarter’ receiving people with pre-existing lower FEV1 during transition from paediatric centres. In the UK, people with CF typically transition to adult centres at the age of 16 years. There were no clear differences in FEV1 between the three groups of centres among those aged 16–17 years. In particular, FEV1 for centres in the ‘lower quarter’ were very similar to centres in the ‘middle half’ in this transition age group. However, differences in FEV1 were apparent among those aged 18–19 years (see table 1E below), which remained throughout all the other age ranges (see Table 2 of the main text).

Table 1E: Annual review %FEV1 among the cohort of people aged 16–19 years for the three groups of specialist CF centres for 2013-2014

|  |  |  |
| --- | --- | --- |
| % predicted FEV1 at annual review, median (IQR) | 2013 | 2014 |
| Upper quarter | Middle half | Lower quarter | Upper quarter | Middle half | Lower quarter |
| Age 16 – 17 yearsAge 18 – 19 years | 83.7 (77.3 – 97.0)(n = 27) 79.2 (61.1 – 90.3)(n = 90) | 72.8 (54.3 – 91.3)(n = 37) 75.9 (54.5 – 91.5)(n = 170) | 73.0 (42.0 – 92.8)(n = 25) 72.2 (56.5 – 82.7)(n = 108) | 80.0 (64.9 – 97.8)(n = 18) 80.4 (61.2 – 92.0)(n = 88) | 79.0 (58.5 – 94.1)(n = 44) 80.3 (60.1 – 94.3)(n = 211) | 78.0 (69.0 – 104.1)(n = 16) 70.8 (54.1 – 86.9)(n = 111) |

\* Only descriptive statistics provided because the small sample size preclude null hypothesis significance testing.

A sensitivity analysis using best FEV1 data for 2014 showed the same trend, whereby centres in upper quarter have the highest age-for-age best annual % FEV1 although this was not the case for the transition age group (see table 2E below). This provides further evidence of the genuine differences in outcomes across the three groups of CF centres. This sensitivity analysis also demonstrates that potential bias in the annual review FEV1 data were balanced by the aggregation of individual CF centres into three different groups (see Appendix C for detailed explanation regarding the analysis method used). Best annual FEV1 data were only available for 924/4269 (21.6%) adults in 2013, hence we did not repeat the same analysis using the 2013 best annual FEV1 data.

Table 2E: Best annual % predicted FEV1 for the three groups of specialist CF centres for 2014, stratified according to age:

|  |  |
| --- | --- |
| Best annual % predicted FEV1, median (IQR) | 2014 |
| Upper quarter(n = 614) | Middle half(n = 1772) | Lower quarter(n = 609) |
| Age 16 – 17 yearsAge 18 – 19 yearsAge 20 – 21 yearsAge 22 – 23 yearsAge 24 – 25 yearsAge 26 – 27 yearsAge 28 – 30 yearsAge 31 – 33 yearsAge 34 – 37 yearsAge 38 – 44 yearsAge ≥45 years | 89.3 (67.5 – 104.6)85.8 (66.9 – 98.9)81.3 (66.7 – 95.7)83.3 (59.6 – 94.3)79.5 (69.9 – 91.1)66.9 (46.3 – 84.2)76.3 (61.7 – 92.0)67.1 (44.8 – 82.8)67.3 (50.9 – 87.4)76.7 (59.8 – 94.7)72.1 (53.5 – 94.7) | 93.2 (67.0 – 105.6)84.1 (68.1 – 97.4)82.5 (63.2 – 99.4)71.7 (56.6 – 91.5)73.7 (55.5 – 91.9)63.6 (45.5 – 83.1)69.1 (51.5 – 87.4)65.1 (47.4 – 86.4)68.0 (52.2 – 84.0)67.2 (52.2 – 87.3)64.7 (46.1 – 88.5) | 105.6 (58.8 – 121.8)74.3 (58.6 – 87.9)80.4 (51.4 – 93.8)74.8 (54.7 – 89.3)78.1 (53.7 – 89.7)59.2 (48.1 – 83.5)69.1 (40.9 – 82.5)61.3 (40.8 – 81.1)58.6 (43.8 – 78.9)65.6 (43.7 – 82.6)56.0 (35.4 – 72.8) |

 van Elteren test p-value (with Bonferroni correction)

**Upper quartile vs lower quartile < 0.001**

**Middle half vs lower quartile < 0.001**

**Upper quartile vs middle half 0.007**

Another evidence supporting robustness of the ranking process is the stepwise increase for both unadjusted %FEV1 and BMI from ‘lower quarter’ to ‘upper quarter’ (see Table 1 of the main text). This mirrored the results from the ESCF analyses using the 1995-1996 dataset [1]. Higher BMI is associated with higher FEV1 [2]; but whereas FEV1 declines with age [3], BMI increases with age among adults with CF [4]. Graphs displaying these relationships with the UK CF registry data are shown in Figure 1E below. The observed stepwise relationship therefore provides further reassurance that the ranking process is robust in identifying centres with genuinely better outcomes instead of just favouring centres with younger adults.

Figure 1E: The relationships between % FEV1 & BMI, % FEV1 & age and BMI & age

 Scatter plot of % FEV1 vs BMI in 2014 Scatter plot of % FEV1 vs age in 2014 Scatter plot of BMI vs age in 2014

% FEV1 is positively correlated with BMI % FEV1 is negatively correlated with age BMI is positively correlated with age

\*The Local Polynomial Regression (LOESS) curve is a non-parametric method for fitting smooth curves to empirical data, to depict relationships between variables [5].

REFERENCES:

1. Johnson C, Butler SM, Konstan MW, et al. Factors influencing outcomes in cystic fibrosis: a center-based analysis. *Chest* 2003;123:20-7.
2. Stallings VA, Stark LJ, Robinson KA, et al. Evidence-based practice recommendations for nutrition-related management of children and adults with cystic fibrosis and pancreatic insufficiency: results of a systematic review. *J Am Diet Assoc* 2008;108:832-9.
3. Liou TG, Elkin EP, Pasta DJ, et al. Year-to-year changes in lung function in individuals with cystic fibrosis. *J Cyst Fibros* 2010;9:250-6.
4. Boelle PY, Viviani L, Busson PF, et al. Reference percentiles for FEV(1) and BMI in European children and adults with cystic fibrosis. *Orphanet J Rare Dis* 2012;7:64.
5. Cleveland WS, Devlin SJ. Locally weighted regression: an approach to regression analysis by local fitting. *J Am Stat Assoc* 1988; 83: 596-610.

**Appendix F: Exploration of case-mix as potential confounding for differences in IV use between the three different groups of CF centres**

Controlling for confounding factor is crucial to expose true relationships between variables and to avoid finding spurious relationships. Tables 3 and 4 in the main text show similar IV use across the different groups of CF centres in 2014, but centres in the ‘lower quarter’ have lower FEV1 and people with lower FEV1 tend to have more exacerbations (i.e. higher IV use). By stratifying IV days according to FEV1, it is clear that centres in the ‘upper quarter’ and ‘middle half’ have higher IV days compared to centres in the ‘lower quarter’.

Stratifying the study subjects according to FEV1 also dealt with other case-mix factors that could influence IV days, including *P. aeruginosa* status, pancreatic status, gender and BMI as shown in the tables below.

*P. aeruginosa* status for the three groups of specialist CF centres ('upper quarter', 'middle half', 'lower quarter') for 2013-2014, stratified according to FEV1:

|  |  |  |
| --- | --- | --- |
| People with chronic *P. aeruginosa* infection (%) | 2013 | 2014 |
| Upper quarter(n = 981) | Middle half(n = 1867) | Lower quarter(n = 1244) | Upper quarter(n = 871) | Middle half(n = 2167) | Lower quarter(n = 1411) |
| FEV1 < 40%FEV1 40% to 69.9%FEV1 ≥ 70% | 113 (83.7)228 (61.3)168 (35.3) | 237 (69.1)453 (63.2)312 (38.6) | 220 (74.8)286 (61.6)168 (34.6) | 90 (72.0)175 (58.5)129 (28.9) | 240 (66.9)478 (58.4)358 (36.2) | 252 (76.1)340 (63.1)198 (36.6) |

 Cochran–Mantel–Haenszel test p-value (with Bonferroni correction)

 2013 2014

Upper quarter vs lower quarter 1.000 **0.013**

Middle half vs lower quarter 1.000 0.062

Upper quarter vs middle half 1.000 0.466

Pancreatic status for the three groups of specialist CF centres ('upper quarter', 'middle half', 'lower quarter') for 2013-2014, stratified according to FEV1:

|  |  |  |
| --- | --- | --- |
| People with pancreatic insufficiency\* (%) | 2013 | 2014 |
| Upper quarter(n = 972) | Middle half(n = 1835) | Lower quarter(n = 1240) | Upper quarter(n = 871) | Middle half(n = 2167) | Lower quarter(n = 1411) |
| FEV1 < 40%FEV1 40% to 69.9%FEV1 ≥ 70% | 125 (95.4)308 (84.2)359 (75.6) | 301 (89.1)605 (85.8)604 (76.3) | 266 (90.8)401 (86.8)359 (74.0) | 116 (92.8)243 (81.3)332 (74.6) | 322 (90.2)698 (86.8)753 (76.7) | 294 (89.8)474 (88.4)396 (73.2) |

\* Data for pancreatic status missing among 45 adults in 2013 and 29 adults in 2014.

 Cochran–Mantel–Haenszel test p-value (with Bonferroni correction)

 2013 2014

Upper quarter vs lower quarter 1.000 1.000

Middle half vs lower quarter 1.000 1.000

Upper quarter vs middle half 1.000 0.310

Gender for the three groups of specialist CF centres ('upper quarter', 'middle half', 'lower quarter') for 2013-2014, stratified according to FEV1:

|  |  |  |
| --- | --- | --- |
| Females (%) | 2013 | 2014 |
| Upper quarter(n = 981) | Middle half(n = 1867) | Lower quarter(n = 1244) | Upper quarter(n = 871) | Middle half(n = 2167) | Lower quarter(n = 1411) |
| FEV1 < 40%FEV1 40% to 69.9%FEV1 ≥ 70% | 56 (41.5)174 (46.8)219 (46.0) | 158 (46.1)351 (49.0)362 (44.7) | 117 (39.8)224 (48.3)207 (42.6) | 58 (46.4)147 (49.2)200 (44.7) | 158 (44.0)382 (46.6)443 (44.8) | 150 (45.3)254 (47.1)226 (41.8) |

 Cochran–Mantel–Haenszel test p-value (with Bonferroni correction)

 2013 2014

Upper quarter vs lower quarter 1.000 0.922

Middle half vs lower quarter 0.564 1.000

Upper quarter vs middle half 1.000 1.000

BMI for the three groups of specialist CF centres ('upper quarter', 'middle half', 'lower quarter') for 2013-2014, stratified according to FEV1:

|  |  |  |
| --- | --- | --- |
| BMI, median (IQR) | 2013 | 2014 |
| Upper quarter(n = 978) | Middle half(n = 1863) | Lower quarter(n = 1229) | Upper quarter(n = 866) | Middle half(n = 2149) | Lower quarter(n = 1410) |
| FEV1 < 40%FEV1 40% to 69.9%FEV1 ≥ 70% | 21.0 (18.7 – 23.4)21.8 (19.9 – 24.0)23.2 (21.1 – 25.7) | 20.0 (18.0 – 22.6)22.0 (19.9 – 24.0)22.9 (20.9 – 25.6) | 20.2 (18.6 – 22.7)21.6 (19.7 – 24.2)23.1 (21.2 – 25.3) | 20.9 (18.6 – 23.5)21.5 (19.6 – 23.3)23.0 (21.0 – 25.5) | 20.2 (18.3 – 22.6)21.9 (20.0 – 24.3)23.2 (21.2 – 25.7) | 20.1 (18.5 – 22.5)21.9 (20.0 – 24.1)23.1 (20.9 – 25.4) |

 van Elteren test p-value (with Bonferroni correction)

 2013 2014

Upper quarter vs lower quarter 0.181 1.000

Middle half vs lower quarter 1.000 1.000

Upper quarter vs middle half 0.208 0.775

Centres in the ‘lower quarter’ have slightly older adults compared to centres in the ‘middle half’ and ‘upper quarter’.

Age of adults for the three groups of specialist CF centres ('upper quarter', 'middle half', 'lower quarter') for 2013-2014, stratified according to FEV1:

|  |  |  |
| --- | --- | --- |
| Age in years, median (IQR) | 2013 | 2014 |
| Upper quarter(n = 981) | Middle half(n = 1867) | Lower quarter(n = 1244) | Upper quarter(n = 871) | Middle half(n = 2167) | Lower quarter(n = 1411) |
| FEV1 < 40%FEV1 40% to 69.9%FEV1 ≥ 70% | 29 (24 – 39)27 (23 – 34)25 (21 – 34) | 28 (23 – 35)28 (23 – 36)26 (21 – 33) | 31 (25 – 39)30 (23 – 38)27 (21 – 34) | 30 (24 – 38)28 (24 – 35)26 (21 – 33) | 30 (24 – 37)28 (23 – 36)26 (21 – 34) | 32 (25 – 41)31 (24 – 38)27 (22 – 35) |

 van Elteren test p-value (with Bonferroni correction)

 2013 2014

Upper quarter vs lower quarter **0.011** **0.007**

Middle half vs lower quarter **0.001** **< 0.001**

Upper quarter vs middle half 1.000 1.000

However, the small age differences do not explain the differences in IV days. Increasing age was associated with lower FEV1 (see Appendix E), and is thus expected to increase IV days. Despite the slightly higher age of the CF centres in the ‘lower quarter’, IV days were still lower. The differences in IV days were therefore not due to differences in case-mix between the centres. Between-group differences in IV use also persisted even after regression modelling that accounted for case-mix factors including age (see Appendix B).

Of note, the difference in age between groups also does not imply that centres in ‘lower quarter’ have better survival or better outcomes. Full explanation of this is provided in Appendix E.

We have also previously explored the centre-by-centre differences in IV use with a stepwise binary logistic regression model that accounts for all relevant case-mix factors [1]. After prior-year IV use and FEV1, the CF centre which an adult with CF attended was the third strongest predictor for the amount of IV antibiotics prescribed for an individual [1]. Therefore, there are genuine differences in IV use across the different adult CF centres in the UK that cannot be explained by differences in case-mix factors.

REFERENCES:

1. Hoo ZH, Wildman MJ, Curley R, Walters SJ, Campbell MJ. Rescue therapy within the UK CF registry: an exploration of the predictors of intravenous antibiotic use amongst adults with cystic fibrosis. *Respirology* 2017 [in press]

**Appendix G: Further analyses of the differences in preventive therapies prescription between the three different groups of CF centres**

The ECFS study using 1995-1996 dataset found higher inhaled antibiotics prescription rates among children and higher dornase alfa prescription rates among adults in CF centres with the best FEV1. It is therefore important to consider whether the use of specific preventive therapies differs across the three groups of CF centres in this study. In this appendix, we explored in detail the between-group differences for each type of preventive therapies (inhaled antibiotics, inhaled mucolytics, long-term oral macrolide) using stratification method to mirror the 1995-1996 ESCF analysis. Methods and the results for regression modelling of preventive therapies data are presented in Appendix B.

The UK CF registry does not collect objective adherence data, or even medication possession ratio (MPR) data as a surrogate marker for adherence. The only data available is the “prescription data”. Inhaled therapy data are collected by the UK CF registry with a check-box for the various medications (e.g. colistin, tobramycin solution, dornase alfa, etc.). Long-term oral macrolide data are collected by the UK CF registry with a check-box for “chronic macrolide”. Therefore, if a check-box is not selected, it is assumed the person is not on that particular treatment. It is however difficult to distinguish missing data from “not on a treatment”.

The following analyses of the prescription data assumed no missing data. A person is considered to have been prescribed anti-Pseudomonas inhaled antibiotic if at least one of the following treatments was selected: colistin, tobramycin solution, colistimethate, tobramycin dry powder, aztreonam, meropenem, ceftazidime and promixin®. A person is considered to have been prescribed inhaled mucolytic if at least one of the following treatments was selected: hypertonic saline and dornase alfa. A person is considered to have been prescribed long-term oral macrolide if the “chronic macrolide” check-box is selected.

There was no clear signal in the prescription of inhaled antibiotics among the three different groups of CF centres. Following stratification with FEV1 to deal with differences in case-mix, centres in the ‘upper quarter’ appeared to be prescribing slightly more inhaled antibiotics while centres in the ‘middle half’ appeared to be prescribing the least inhaled antibiotics.

Prescription of inhaled anti-Pseudomonal antibiotics across the three groups of specialist CF centres ('upper quarter', 'middle half', 'lower quarter') for 2013-2014:

|  |  |  |
| --- | --- | --- |
| People prescribed inhaled antibiotics (%) | 2013 | 2014 |
| Upper quarter(n = 983) | Middle half(n = 1869) | Lower quarter(n = 1244) | Upper quarter(n = 871) | Middle half(n = 2167) | Lower quarter(n = 1411) |
| Overall | 677 (68.9) | 1157 (61.9) | 747 (60.0) | 596 (68.4) | 1475 (68.1) | 1052 (74.6) |
| FEV1 < 40%FEV1 40% to 69.9%FEV1 ≥ 70% | 123 (91.1)286 (76.9)268 (56.3) | 256 (74.6)500 (69.7)401 (49.6) | 222 (75.5)297 (64.0)228 (46.9) | 112 (89.6)241 (80.6)243 (54.4) | 291 (81.1)610 (74.5)574 (58.0) | 299 (90.3)435 (80.7)318 (58.3) |

 Cochran–Mantel–Haenszel test p-value (with Bonferroni correction)

 2013 2014

Upper quarter vs lower quarter **< 0.001** 0.779

Middle half vs lower quarter 0.254 **0.008**

Upper quarter vs middle half **< 0.001** 1.000

There were clearer signals in the prescription of inhaled mucolytics. The direction of difference is similar to that of inhaled antibiotics, in that centres in the ‘upper quarter’ appeared to be prescribing the most inhaled mucolytics while centres in the ‘middle half’ appeared to be prescribing the least inhaled mucolytics.

Prescription of inhaled mucolytics across the three groups of specialist CF centres ('upper quarter', 'middle half', 'lower quarter') for 2013-2014:

|  |  |  |
| --- | --- | --- |
| People prescribed inhaled mucolytics (%) | 2013 | 2014 |
| Upper quarter(n = 983) | Middle half(n = 1869) | Lower quarter(n = 1244) | Upper quarter(n = 871) | Middle half(n = 2167) | Lower quarter(n = 1411) |
| Overall | 804 (81.8) | 1160 (62.1) | 862 (69.3) | 705 (80.9) | 1442 (66.5) | 1048 (74.3) |
| FEV1 < 40%FEV1 40% to 69.9%FEV1 ≥ 70% | 125 (92.6)333 (89.5)346 (72.7) | 278 (81.0)497 (69.3)385 (47.6) | 254 (86.4)358 (77.2)250 (51.4) | 115 (92.0)266 (89.0)324 (72.5) | 298 (83.0)604 (73.7)540 (54.6) | 289 (87.3)433 (80.3)326 (60.3) |

 Cochran–Mantel–Haenszel test p-value (with Bonferroni correction)

 2013 2014

Upper quarter vs lower quarter **< 0.001** **< 0.001**

Middle half vs lower quarter **0.002** **< 0.001**

Upper quarter vs middle half **< 0.001** **< 0.001**

Differences in the prescription of long-term oral macrolide were inconsistent between 2013 and 2014. In 2013, centres in the ‘lower quarter’ have the lowest prescription of long-term oral macrolide, but the prescription rates were similar between all three groups of CF centres in 2014.

Prescription of long-term oral macrolide across the three groups of specialist CF centres ('upper quarter', 'middle half', 'lower quarter') for 2013-2014:

|  |  |  |
| --- | --- | --- |
| People prescribed long-term oral macrolide (%) | 2013 | 2014 |
| Upper quarter(n = 983) | Middle half(n = 1869) | Lower quarter(n = 1244) | Upper quarter(n = 871) | Middle half(n = 2167) | Lower quarter(n = 1411) |
| Overall | 593 (60.3) | 1237 (66.2) | 697 (56.0) | 496 (56.9) | 1300 (60.0) | 852 (60.4) |
| FEV1 < 40%FEV1 40% to 69.9%FEV1 ≥ 70% | 107 (79.3)256 (68.8)230 (48.3) | 280 (81.6)499 (69.6)458 (56.6) | 213 (72.4)299 (64.4)185 (38.1) | 101 (80.8)191 (63.9)204 (45.6) | 264 (73.5)546 (66.7)490 (49.5) | 242 (73.1)357 (66.2)253 (46.8) |

 Cochran–Mantel–Haenszel test p-value (with Bonferroni correction)

 2013 2014

Upper quarter vs lower quarter **0.001**  1.000

Middle half vs lower quarter **< 0.001** 1.000

Upper quarter vs middle half 0.493 1.000

It is unlikely that differences in the prescription of preventive therapies alone could explain the differences in FEV1 between the three groups of CF centres, since the differences were not consistent for 2013-2014, unlike IV use. For example, centres in the lower quarter prescribed less long-term oral macrolide in 2013, but this was not the case in 2014. The direction of differences is also somewhat paradoxical. Higher prescription of preventive inhaled therapy should not improve FEV1 of the ‘upper quarter’ (compared to the ‘middle half’ in 2013) while at the same time reduce FEV1 of the ‘lower quarter’ (compared to the ‘middle half’ in 2014). A possible explanation is that inadequate use of IV antibiotics has such detrimental effect on FEV1 that could not be compensated with slightly higher prescription of preventive inhaled therapy.

However, limitations of the preventive therapy data should be considered in the interpretation of these results. Based on the current data collection system, a person being prescribed a month’s supply of dornase alfa that he / she did not use at all is considered to have achieved the same outcome as someone who has been 100% adherent to 12-months of dornase alfa prescription (i.e. check-box for dornase alfa selected). Yet the potential benefit derived from dornase alfa for these two theoretical adults with CF would be vastly different. Unfortunately, the massive inter-individual variations in the theoretical examples provided are common occurrences in the real world [1]. There are also potential between-centre differences in adherence to inhaled therapies [2]. A study comparing the MPR of tobramycin inhalation powder and tobramycin inhalation solution did not find any differences with univariate analysis or multivariable analysis to account for clinical characteristics [2]. Differences in MPR were only detected following adjustment for geographical location (which was a proxy for which CF centre a person attended) [2].

Previous studies also showed significant disconnect between medications prescribed and medications that are actually collected, with MPR of around 50% for preventive inhaled therapies [3]. Not all prescriptions that are collected will eventually be used correctly. There is often a discrepancy between MPR and objective adherence data of around 20%, even for simple and very potent treatments such as oral ivacaftor [4,5]. Objective adherence measurement in adults suggests that median adherence is less than 36% [1], also emphasising the limitation of inhaled therapy prescription data. Preventive therapy data was perhaps not recorded as reliably as IV antibiotics within the UK CF registry. For example, it seems unlikely that inhaled antibiotic prescriptions among people with FEV1 <40% at the ‘lower quarter’ centres would increase from 75.5% in 2013 to 90.3% in 2014. Inaccurate data could potentially introduce bias and produce spurious relationships. It is also possible that some CF centres may just be more vigilant in recording their prescriptions instead of genuinely utilising preventive therapies more effectively.

Therefore, these preventive therapy prescription results should be interpreted with caution. We propose that a platform that automatically downloads objective adherence data via electronic data capture should be developed to study the real world effects of preventive inhaled therapies. However, the limitations of preventive therapy prescription data should not diminish the significance of the IV use findings. The very strong and consistent signals from the IV use analysis using both regression modelling and stratification method would suggest that differences in IV use are much more likely to influence the FEV1 outcomes among the UK adult specialist CF centres, compared to prescription of preventive therapies.

REFERENCES:

1. Daniels T, Goodacre L, Sutton C, Pollard K, Conway S, Peckham D. Accurate assessment of adherence: self-report and clinician report vs electronic monitoring of nebulizers. *Chest* 2011;140:425-32.
2. Hamed K, Conti V, Tian H, Loefroth E. Adherence to tobramycin inhaled powder vs inhaled solution in patients with cystic fibrosis: analysis of US insurance claims data. Patient Prefer Adherence 2017;11:831-38.
3. Quittner AL, Zhang J, Marynchenko M, et al. Pulmonary medication adherence and health-care use in cystic fibrosis. *Chest* 2014;146:142-51.
4. Modi AC, Lim CS, Yu N, et al. A multi-method assessment of treatment adherence for children with cystic fibrosis. *J Cyst Fibros* 2006;5:177-85.
5. Siracusa CM, Ryan J, Burns L, et al. Electronic monitoring reveals highly variable adherence patterns in patients prescribed ivacaftor. *J Cyst Fibros* 2015;14:621-6.