Article:
Hoo, Z. orcid.org/0000-0002-7067-3783, Campbell, M.J., Walters, S.J. et al. (1 more author) (2019) Understanding FEV1 for the purpose of cystic fibrosis registry comparisons: Does bias in annual review FEV1 affect between-centre comparison within the UK? An analysis of registry data. Journal of Evaluation in Clinical Practice. ISSN 1356-1294

https://doi.org/10.1111/jep.13097

Title: Understanding FEV_1 for the purpose of cystic fibrosis registry comparisons: does bias in annual review FEV_1 affect between-centre comparison within the UK? An analysis of registry data.

Short title: Data issues when comparing CF registry outcomes

Author names (surnames are underlined) and affiliations:

Zhe Hui Hoo[1,2] z.hoo@sheffield.ac.uk
Room 1.03, Innovation Centre, 217, Portobello, Sheffield S1 4DP, United Kingdom.

Michael J. Campbell[1] m.j.campbell@sheffield.ac.uk
ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, United Kingdom

Stephen J. Walters[1] s.j.walters@sheffield.ac.uk
ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, United Kingdom

Martin J. Wildman[2,1] martin.wildman@sth.nhs.uk
Sheffield Adult Cystic Fibrosis Centre, Brearley Outpatient, Northern General Hospital, Herries Road, Sheffield S5 7AU, United Kingdom.

[1] School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK
[2] Sheffield Adult CF Centre, Northern General Hospital, Sheffield, UK

Corresponding author:
Zhe Hui Hoo
Room 1.03, Innovation Centre, 217, Portobello, Sheffield S1 4DP, United Kingdom.
Email: z.hoo@sheffield.ac.uk Telephone: 0114 226 6281

ABSTRACT

Rationale, aims and objective:

We previously demonstrated that annual review %FEV_1 under-estimates lung health of adults with CF compared to %FEV_1 captured during periods of clinical stability. This has implications in the comparisons against registries with encounter-based FEV_1, such as the US. It is uncertain whether this bias affects between-centre comparison within the UK. Previous funnel plot analyses have identified variation in annual review %FEV_1 according to centre size, hence we investigated whether paired differences between annual review and best %FEV_1 also vary according to centre size.

Methods:

This registry analysis included 18 adult CF centres in the UK with ≥80% completeness for best FEV_1 data in 2014. Mean discrepancy between annual review and best %FEV_1 is a surrogate for the extent by which annual review %FEV_1 underestimates lung health; and was plotted against centre size. A Local Polynomial Regression (LOESS) curve was used to explore the relationship between the two variables. An appropriate model is fitted based on the LOESS curve to determine the strength of relationship between discrepancies in %FEV_1 and centre size.

Results:

There is an inverted U-shaped relationship between mean discrepancies in %FEV_1 and centre size. A regression of the paired mean difference in %FEV_1 against centre size showed a significant improvement in the goodness of fit for a quadratic model (R^2 = 23.8% for a quadratic model compared with 0.4% for a linear one; p = 0.048 for the quadratic term).

Conclusions:

Annual review %FEV_1 under-estimated lung health of adults from small and large centres in the UK to a greater extent compared to medium-sized centres. A plot of %FEV_1 against centre size (e.g. funnel plot comparison) would be affected by systematic bias in annual review %FEV_1. Therefore, annual review %FEV_1 is an unreliable metric to compare health outcomes of adult CF centres within the UK.

KEYWORDS

Cystic fibrosis
Clinical epidemiology
Registry analysis
Respiratory function tests

INTRODUCTION

Cystic fibrosis (CF) is an archetypal long-term health condition – it is progressive and life-limiting with no cure at present, but an array of effective medications is available to improve prognosis. Median CF survival is currently 45-50 years, compared to survival of around 6 months in the 1940's.[1,2] This dramatic improvement is partly due to increasing numbers of efficacious therapeutic options.[3] However, improvement in the quality of care provided by specialist CF centres is also crucial.[4] Indeed, quality improvement initiatives e.g. centre-comparisons and benchmarking have been critical in transforming healthcare delivery in CF.[5-8]

The UK CF registry annual data report plots centre-level $\%FEV_1$ as a ranked graph (caterpillar plot) and also as a funnel plot, with the aim of using data to drive quality improvement.[9] In a funnel plot, the outcome of interest is plotted against centre size with control limits narrowing as the centre size increases.[10] Centres that lie outside the control limits of a funnel plot arguably have outcomes that reflect a genuine difference from the average, i.e. a funnel plot implies that centres lying outside the control limits are 'better' or 'worse' than expected.[10]

Annual review $\%FEV_1$ has traditionally been used as the $\%FEV_1$ metric for comparing centres in the UK CF registry annual data report.[11] Recently a greater emphasis has been placed on best FEV_1 and this allows us to explore the differences between the two metrics. Previous funnel plot analyses have identified an inverted U-shape relationship between annual review $\%FEV_1$ and centre size, with negative outliers tended to be smaller (<100 adults) and larger (>350 adults) centres instead of medium-sized centres.[12] Taken at face value, this might imply that medium-sized adult CF centres in the UK achieved the best health outcomes and standards of care.

However, the 'pyramid of investigation' model emphasises that data must first be sufficiently robust before drawing other conclusions (e.g. regarding quality of care) in observational epidemiology.[13] Our previous analysis demonstrated that annual review $\%FEV_1$ underestimates lung health of adults with CF in the UK in comparison to $\%FEV_1$ captured during periods of clinical stability.[14] It is uncertain whether the bias in annual review $\%FEV_1$ affects between-centre comparison within the UK. It may be that annual review $\%FEV_1$ under-estimates centre-level lung health for all adult CF centre by the same extent because procedures pertaining to annual review could be identical in all centres. Alternatively, the bias in annual review $\%FEV_1$ may differ between different centres such that between-centre comparisons within the UK are also affected.

We investigated this issue using data from the UK CF registry to determine whether discrepancies between annual review $\%FEV_1$ and best annual $\%FEV_1$ vary according to centre size and thus affect between-centre comparisons using a funnel plot analysis.

METHODS AND MATERIALS

This is a cross-sectional analysis using the UK CF registry data for 2014 from 18 UK adult CF centres with ≥80% completeness for best FEV_1 data. Ten centres with <80% completeness for best FEV_1 data were excluded because comparison between annual review and best $\%FEV_1$ are less reliable with large amount of missing best FEV_1 data. Adults who had a lung transplantation were also excluded because lung transplantation has transformative effects on lung health,[15] such that their FEV_1 no longer represent that of a typical adult with CF. NHS research ethics approval (Huntingdon Research Ethics Committee 07/Q0104/2) was granted for the UK CF Registry. Under the terms of the NHS ethics approval, the UK CF Trust steering committee approved this study.

The UK CF registry data were collected during annual reviews between 1st January and 31st December 2014. Data for age, gender, pancreatic status, CF related diabetes and body mass index (BMI, in kg/m^2) were obtained to understand the demographics of the participants included in this analysis. Annual review FEV_1 data were the FEV_1 readings taken during annual review. In addition to annual review FEV_1 data, best FEV_1 data have been collected by the UK CF registry since 2012 for the European CF Society Patient registry.[16] The best FEV_1 data represent the highest FEV_1 reading in the 1-year period prior to the date of annual review (i.e. if a person had annual review on 1st July 2014, the highest FEV_1 reading between 1st July 2013 to 1st July 2014 should be that person's 'best FEV_1' for 2014). Both the annual review and best $\%FEV_1$ were calculated using Knudson equation.[17]

The magnitude of group-level mean discrepancy between paired annual review and best $\%FEV_1$ is a surrogate for the extent by which annual review $\%FEV_1$ underestimates lung health of adults with CF – a larger discrepancy indicates greater underestimation of centre-level lung health and vice versa.[14] Therefore, centre-by-centre discrepancy between paired annual review and best $\%FEV_1$ were calculated, then plotted against centre size. A Local Polynomial Regression (LOESS) curve was used to explore the relationship between centre-level mean discrepancy in $\%FEV_1$ and centre size. The LOESS curve was generated using default SPSS v24.0 setting (Kernel: Epanechnikov, 50% of points to fit). LOESS curve is a non-parametric method for fitting smooth curves to data, to depict relationships between variables.[18] An appropriate parametric model was then fitted based on the relationship depicted by the LOESS curve, to determine the strength of relationship between centre-level discrepancies in $\%FEV_1$ vs centre size. Centre-level annual review $\%FEV_1$ adjusted for age and pancreatic status (to account for differences in case-mix between centres) was also plotted on a funnel plot for all 28 adult CF centres. This allows identification of positive and negative outliers, elucidation of the relationship between centre-level annual review $\%FEV_1$ vs centre size, and comparison with the plot of centre-level discrepancies in $\%FEV_1$ vs centre size.

The following sensitivity analyses to further understand the relationship between centre-level discrepancies in $\%FEV_1$ vs centre size were also reported in the online appendix: LOESS curve fitted through 25% and 75% of points, LOESS curve and the parametric model fitted after excluding the

extreme left and extreme right points on the scatterplot, weighted analysis for the parametric model, analyses repeated using paired median difference in $\%FEV_1$ (instead of mean difference) and analyses repeated with centre size truncated at 150 (i.e. around 150 adults were randomly sampled in centres with more than 150 adults). Analyses were performed using SPSS v24 (IBM Corp, Armonk, NY, USA), Stata 14 (StatCorp, College Station, TX, USA) and R v3.3.0 (R Foundation for Statistical Computing, Vienna, Austria). A *p*-value <0.05 was considered to be statistically significant.

RESULTS

In total, there were 4925 adults without lung transplantation across 28 adult CF centres in the UK. Of these, 3149 (63.9%) were from the 18 centres with ≥80% completeness for best FEV_1 data in 2014. Of the 15 medium-sized centres (100–350 adults), 12 (80%) had ≥80% completeness for best FEV_1 data, whereas this level was only achieved by 5 (50%) of the 10 smaller centres (<100 adults) and by 1 (33%) of the 3 larger centres (>350 adults). Adults from centres with either ≥80% or <80% completeness for best FEV_1 data were similar in terms of gender, pancreatic status and CF related diabetes (see Table 1). Adults from centres with ≥80% completeness for best FEV_1 data were slightly younger (median age 27 years vs 29 years), which may partly explain the slightly higher annual review $\%FEV_1$ for that group (median 67.6% vs 64.4%).

After adjusting for age and pancreatic status using a generalised linear model (GLM) approach as previously described,[19] a funnel plot analysis for all 28 adult CF centres in the UK with annual review $\%FEV_1$ as the outcome of interest identified six centres (four of which had <100 adults, two of which had >350 adults) as negative outliers (i.e. lower than expected $\%FEV_1$) and one centre with 100–350 adults as a positive outlier (i.e. higher than expected $\%FEV_1$), see Figure 1a. The LOESS curve of $\%FEV_1$ versus centre size showed an inverted U-shaped relationship between case-mix adjusted annual review $\%FEV_1$ and centre size (Figure 1b). A regression of the $\%FEV_1$ against centre size showed a significant improvement in the goodness of fit for a quadratic model (R^2 = 28.8% for a quadratic model compared with 10.5% for a linear one; p = 0.018 for the quadratic term). The quadratic model is $y = b_0 + b_1x + b_2x^2$ where y = case-mix adjusted annual review $\%FEV_1$, x = number of adults in each centre.

Among the 18 centres with ≥80% completeness for best FEV_1 data, there was also an inverted U-shaped relationship between centre-level paired mean difference in $\%FEV_1$ and centre size (see Figure 2). A regression of the paired mean difference in $\%FEV_1$ against centre size showed a significant improvement in the goodness of fit for a quadratic model (R^2 = 23.8% for a quadratic model compared with 0.4% for a linear one; p = 0.048 for the quadratic term). After excluding the extreme left and extreme right points on the scatterplot, similar results were obtained (R^2 = 30.0% for a quadratic model compared with 0.7% for a linear one; p = 0.036 for the quadratic term), see Appendix B. Broadly similar results were also obtained with other sensitivity analyses detailed in Appendices A-E.

Given that the magnitude of group-level mean discrepancy between paired annual review and best $\%FEV_1$ is a surrogate for the extent by which annual review $\%FEV_1$ underestimates lung health, this finding suggests that annual review $\%FEV_1$ underestimates lung health to a greater extent in smaller and larger centres. Therefore, the relationship observed in a funnel plot analysis among adult CF centres within the UK with annual review $\%FEV_1$ as the outcome of interest may well be exaggerated by the systematic bias in annual review FEV_1 data.

DISCUSSION

We previously demonstrated that annual review $\%FEV_1$ under-estimates lung health of adults in comparison to $\%FEV_1$ captured during periods of clinical stability, which is an important systematic bias to consider for cross-country comparisons against registries with encounter-based FEV_1.[14] In this manuscript, we extended our previous work and analysed centre-by-centre discrepancy between paired annual review and best $\%FEV_1$ to determine if systematic bias in annual review FEV_1 also affects between-centre comparisons within the UK. We found a statistically significant quadratic relationship between centre-level paired mean difference in $\%FEV_1$ and centre size. This mirrored the relationship between case-mix adjusted annual review $\%FEV_1$ and centre size in a funnel plot analysis; implying that systematic bias in annual review $\%FEV_1$ according to centre size exaggerates the relationship observed in the funnel plot. As such, annual review $\%FEV_1$ is an unreliable metric to compare health outcomes of adult CF centres within the UK using a funnel plot. Other forms of UK CF registry comparisons using annual review $\%FEV_1$ as the outcome of interest may also be affected by systematic bias.

Our findings emphasise the importance of data issues for accurate and reliable interpretations of results from CF registry comparisons. Registry comparisons are only meaningful if apples are being compared to apples and not to oranges because systematic bias in registry data cannot be easily controlled with statistical methods,[20] even for objective outcomes e.g. survival.[21] Indeed, the 'pyramid of investigation' model advocates an incremental approach to understand variation in health outcomes, starting with review of data before considering other potential sources of variation e.g. differences in case-mix and quality of care.[13] Using annual review $\%FEV_1$ as a metric to infer lung health, it would appear as though adults at medium-sized adult CF centres (i.e. centres with 100–350 adults) have better lung health compared to adults in smaller (<100 adults) or larger (>350 adults) centres in the UK. We suggest that smaller and larger centres may simply have greater difficult in organising annual reviews during periods of clinical stability, hence underestimating the lung health of their population. Certainly for very large centres with a busy annual review schedule (a centre with 600 adults would need to complete an average of more than 11 annual reviews per week), there is less flexibility for re-scheduling another annual review prior to the data entry deadline if an adult were to turn up unwell.

Our findings may also explain some of the results in a recent UK CF registry analysis by Nightingale and Osmond.[9] In that analysis, centre ranking based on annual review %FEV_1 varied substantially from year-to-year. Centre ranking based on annual review %FEV_1 was also substantially different from centre ranking based on change in annual review %FEV_1. Since standard of care provided by an adult CF centre is unlikely to change rapidly from year-to-year, and similar rankings should be achieved with both FEV_1 and change in FEV_1 if both metrics are reliable measures of lung health for adults with CF, Nightingale and Osmond surmised that any apparent FEV_1 difference is unlikely to be related to differing standards of care between centres.[9] In light of our finding that annual review FEV_1 in the UK were not always collected during periods of clinical stability (i.e. lung health is being measured with error by annual review FEV_1), the year-to-year variation in centre ranking and discrepancies between FEV_1 and change in FEV_1 may be related at least in part to the inherent bias of annual review FEV_1. The main driver for FEV_1 fluctuation is pulmonary exacerbation,[22] which is a stochastic event.[23] An adult who turned up during annual review at Year 1 with an exacerbation may be well during annual review at Year 2. The increased variability of annual review FEV_1 from year-to-year (compared to "FEV_1 when stable") means that centre ranking based on annual review FEV_1 would be more prone to fluctuation compared to centre ranking based on "FEV_1 when stable". The increased variability of annual review FEV_1 from year-to-year (compared to "FEV_1 when stable") also means that the calculated change in FEV_1 in an unreliable metric for lung health – a large decline in annual review FEV_1 may simply be due to transient pulmonary exacerbation instead of representing a substantial drop in actual lung health. This would explain the discrepancies in centre ranking based on FEV_1 and change in FEV_1. In other words, it is not that "FEV_1 when stable" is an unreliable metric to infer standards of care for adult CF centres in the UK; but annual review FEV_1 with its inherent bias is an unsuitable metric for centre comparison.

Therefore, it is unnecessary to abandon the use of FEV_1 as a metric for centre comparisons or benchmarking in the UK. However, it is crucial to use a more reliable metric than annual review FEV_1 to accurately reflect lung health. There is currently a two tier data collection system within the UK CF registry. Encounter-based data collection is mandatory for the minority of people with CF that participate in registry-based trials such as CF START (ISRCTN18130649). However, data are only collected from the vast majority of people with CF on an annual basis (during annual reviews). There has been a move by the UK CF registry towards comparisons using best FEV_1 data (which may be more reliable than annual review FEV_1); and it is encouraging that in the 2017 registry annual data report 25/28 (89%) of centres now have ≥80% completeness for best FEV_1 data.[24] Nonetheless, "best FEV_1" reported to the registry could still under-estimate the lung health of people with CF if these data did not reliably select the highest FEV_1 from all FEV_1 readings over a 1-year period.[14] Identification of best FEV_1 data is difficult if data collectors are required to manually extract these data from routine clinical record, particularly in larger centres.

Indeed, best FEV_1 data are most robust if all FEV_1 readings are recorded in a single database via encounter-based data collection, such that the highest reading over a given time period can be automatically and accurately identified. Other than supporting reliable centre comparison using FEV_1 and driving iterative quality improvement cycles, encounter-based data collection for all study subjects (e.g. the US CFFPR[25]) also has other benefits, including allowing robust study of FEV_1 decline trajectory and impact of FEV_1 variability.[22] There are concerns regarding the resource requirements to implement comprehensive encounter-based data collection; but a previous modelling exercise using worst case scenario estimates did suggest that 75 minutes of data entry session per week would be adequate for a medium-sized CF centre to implement this, which is equivalent to a cost of approximately £4 per patient per year.[26] Some adult CF centres in the UK, such as Leeds and Sheffield, have already implemented encounter-based FEV_1 data collection via the use of integrated electronic care records;[27,28] and the CFHealthHub platform (ISRCTN55504164) now in use in 80% of the UK adult CF centres also offers a facility for encounter-based FEV_1 data capture.

The $\%FEV_1$ in this analysis was calculated using Knudson equation to mirror the approach of our previous work.[14] Similar results for the discrepancy between annual review and best $\%FEV_1$ would be obtained with GLI equation because paired difference between two FEV_1 readings was calculated.[29] A limitation of our analysis is data on the number of FEV_1 readings per patient are unavailable due to the lack of encounter-based data entry in the UK CF registry. It may be possible that medium-sized centres were able to review their patients more frequently and to perform more FEV_1 measurements per patient (i.e. have a better chance of identifying the 'true' best FEV_1 reading). Having the data on FEV_1 readings per patient could therefore provide further evidence to support our hypothesis that there is differential data collection by UK adult CF centres according to centre size. It is more likely for larger centres to under-estimate their best FEV_1 compared to medium-sized centres. Thus our findings that larger centres have larger paired FEV_1 differences would be even more obvious if larger centres were not under-estimating their best FEV_1. Another limitation is the extent of missing best FEV_1 data in the UK CF registry. It is also uncertain whether similar findings exist in other years. This analysis was limited to 2014 because best FEV_1 data were only being collected from 2012 onwards, and best FEV_1 data were only available for 780/4380 (17.8%) and 1004/4528 (22.3%) of the adults in 2012 and 2013 respectively. Repeating the analysis using the 2016 and 2017 datasets would help to confirm our findings.

Nonetheless, there is other circumstantial evidence to indicate differential data collection by UK adult CF centres according to centre size. Firstly, medium-sized adult CF centres (i.e. centres with 100–350 adults) were more likely to achieve ≥80% completeness for best FEV_1 data (12/15, 80%) compared to smaller centres (<100 adults; 5/10, 50%) and larger centres (>350 adults; 1/3, 33%). Secondly, centres with <80% completeness for best FEV_1 data were twice as likely to report missing pancreatic status compared to centres with ≥80% completeness for best FEV_1 data (24/1776, 1.4% vs 23/3149, 0.7%).

FEV_1 is often used as an outcome measure for benchmarking exercises in CF[5,8,30,31] since it is an important CF prognostic marker.[32-35] This is the first study to demonstrate empirically the systematic bias in annual review FEV_1 according to CF centre size and its unsuitability as a metric to compare health outcomes of adult CF centres within the UK. In order to reliably discern the quality of CF centres in the UK, the UK CF registry data collection system may require modification.

REFERENCES

1. Burgel PR, Bellis G, Olesen HV, et al. Future trends in cystic fibrosis demography in 34 European countries. *Eur Respir J.* 2015;46(1):133-141.
2. Davis PB. Cystic fibrosis since 1938. *Am J Respir Crit Care Med*. 2006;173(5):475-482.
3. De Boeck K, Amaral MD. Progress in therapies for cystic fibrosis. *Lancet Respir Med.* 2016;4(8):662-674.
4. Mogayzel PJ Jr, Dunitz J, Marrow LC, Hazle LA. Improving chronic care delivery and outcomes: the impact of the cystic fibrosis Care Center Network. *BMJ Qual Saf.* 2014;23(Suppl 1):i3-i8.
5. 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.* 2014;23(Suppl 1):i15-i22.
6. Jackson AD, Goss CH. Epidemiology of CF: How registries can be used to advance our understanding of the CF population. *J Cyst Fibros.* 2018;17(3):297-305.
7. Schechter MS, Fink AK, Homa K, Goss CH. The Cystic Fibrosis Foundation Patient Registry as a tool for use in quality improvement. *BMJ Qual Saf.* 2014;23(Suppl 1):i9-i14.
8. Schechter MS. Benchmarking to improve the quality of cystic fibrosis care. *Curr Opin Pulm Med.* 2012;18(6):596-601.
9. 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;16(5):585-591.
10. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med.* 2005;24 (8):1185-1202.
11. UK Cystic Fibrosis Trust. UK Cystic Fibrosis Registry 2014 Annual Data Report https://www.cysticfibrosis.org.uk/the-work-we-do/uk-cf-registry/reporting-and-resources Accessed September 28, 2018
12. MacNeill SJ, Pierrotti L, Mohammed MA, et al. Quality improvement in CF: what can we learn from each other? A statistical analysis of UK Registry data and consultations with clinicians and patients. *J Cyst Fibros.* 2018;17(Suppl 3):S47-S48.
13. Mohammed MA, Rathbone A, Myers P, Patel D, Onions H, Stevens A. An investigation into general practitioners associated with high patient mortality flagged up through the Shipman inquiry: retrospective analysis of routine data. *BMJ.* 2004;328(7454):1474-1477.

14. Hoo ZH, Curley R, Campbell MJ, Walters SJ, Wildman MJ. The importance of data issues when comparing cystic fibrosis registry outcomes between countries: Are annual review FEV_1 in the UK only collected when subjects are well? *J Eval Clin Pract.* 2018;24(4):745-751.
15. Pego-Fernandes PM, Abrao FC, Fernandes FL, Caramori ML, Samano MN, Jatene FB. Spirometric assessment of lung transplant patients: one year follow-up. *Clinics (Sao Paulo).* 2009;64(6):519-525.
16. Viviani L, Zolin A, Mehta A, Olesen HV. The European Cystic Fibrosis Society Patient Registry: valuable lessons learned on how to sustain a disease registry. *Orphanet J Rare Dis.* 2014;9:81.
17. Knudson RJ, Lebowitz MD, Holberg CJ, Burrows B. Changes in the normal maximal expiratory flow-volume curve with growth and aging. *Am Rev Respir Dis.* 1983;127(6):725-734.
18. Cleveland WS, Devlin SJ. Locally weighted regression: an approach to regression analysis by local fitting. *J Am Stat Assoc.* 1988;83(403):596-610.
19. Hoo ZH, Campbell MJ, Curley R, Walters SJ, Wildman MJ. Do cystic fibrosis centres with the lowest FEV_1 still use the least amount of intravenous antibiotics? A registry-based comparison of intravenous antibiotic use among adult CF centres in the UK. *J Cyst Fibros.* 2018;17(3):360-367.
20. Urschel S. Apples, oranges, and statistical magic: limitations of registry studies and need for collaborative studies. *J Heart Lung Transplant.* 2015;34(9):1136-1138.
21. Sykes J, Stanojevic S, Goss CH, et al. A standardized approach to estimating survival statistics for population-based cystic fibrosis registry cohorts. *J Clin Epidemiol.* 2016;70:206-213.
22. Morgan WJ, VanDevanter DR, Pasta DJ, et al. Forced Expiratory Volume in 1 Second Variability Helps Identify Patients with Cystic Fibrosis at Risk of Greater Loss of Lung Function. *J Pediatr.* 2016;169:116-121.
23. Ferkol T, Rosenfeld M, Milla CE. Cystic fibrosis pulmonary exacerbations. *J Pediatr.* 2006;148(2):259-264.
24. UK Cystic Fibrosis Trust. UK Cystic Fibrosis Registry 2017 Annual Data Report https://www.cysticfibrosis.org.uk/the-work-we-do/uk-cf-registry/reporting-and-resources Accessed September 28, 2018
25. Knapp EA, Fink AK, Goss CH, et al. The Cystic Fibrosis Foundation Patient Registry. Design and Methods of a National Observational Disease Registry. *Ann Am Thorac Soc.* 2016;13(7):1173-1179.
26. Forshaw S, Hoo ZH, Curley R, Wildman MJ. Moving the UK CF registry from annual review data collection to encounter-based data collection: cost estimation to understand feasibility. *J Cyst Fibros.* 2017;16(Suppl 1):S166.
27. Peckham D, Etherington C, White H, et al. The development and deployment of integrated electronic care records in a regional adult and paediatric cystic fibrosis unit. *J Cyst Fibros.* 2014;13(6):681-686.

28. White H, Shaw N, Denman S, Pollard K, Wynne S, Peckham DG. Variation in lung function as a marker of adherence to oral and inhaled medication in cystic fibrosis. *Eur Respir J.* 2017;49(3):1600987.
29. Konstan MW, Wagener JS, VanDevanter DR, et al. Comparison of FEV_1 reference equations for evaluating a cystic fibrosis therapeutic intervention. *Pediatr Pulmonol.* 2017;52(8):1013-1019.
30. Stern M, Niemann N, Wiedemann B, Wenzlaff P; German CFQA Group. Benchmarking improves quality in cystic fibrosis care: a pilot project involving 12 centres. *Int J Qual Health Care.* 2011;23(3):349-356.
31. Wagener JS, Elkin EP, Pasta DJ, et al. Pulmonary function outcomes for assessing cystic fibrosis care. *J Cyst Fibros.* 2015;14(3):376-383.
32. Liou TG, Adler FR, Fitzsimmons SC, et al. Predictive 5-year survivorship model of cystic fibrosis. *Am J Epidemiol.* 2001;153(4):345-352.
33. Corey M, Edwards L, Levison H, Knowles M. Longitudinal analysis of pulmonary function decline in patients with cystic fibrosis. *J Pediatr.* 1997;131(6):809-814.
34. Rosenbluth DB, Wilson K, Ferkol T, Schuster DP. Lung function decline in cystic fibrosis patients and timing for lung transplantation referral. *Chest.* 2004;126(2):412-419.
35. Konstan MW, VanDevanter DR, Sawicki GS, et al. Association of high-dose ibuprofen use, lung function decline, and long-term survival in children with cystic fibrosis. *Ann Am Thorac Soc.* 2018;15(4):485-493.

ACKNOWLEDGEMENTS

We would like to thank the UK CF Trust for supplying the data that made this analysis possible. We would like to acknowledge all the people with CF in the UK who consent to be part of the UK CF registry.

FUNDING

This report is independent research arising from a Doctoral Research Fellowship, Zhe Hui Hoo, DRF-2014-07-092) supported by the National Institute for Health Research. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

COMPETING INTERESTS

MJW is the Chair of the UK CF Registry Research Committee and has argued in favour of shifting the UK CF registry to an encounter-based data entry system. Other co-authors have no conflicts of interest to declare.

Table 1: Characteristics of adults with CF from UK centres with ≥80% and <80% completeness for best FEV_1 data

	2014 UK CF registry data among adult CF centres with ≥80% completeness for best FEV_1 data [§]	2014 UK CF registry data among adult CF centres with <80% completeness for best FEV_1 data [Ω]
Total number of adults, *N*	3149	1776
Number of adults per centres		
Median (IQR)	186 (75 – 241)	111 (53 – 248)
Mean (95% CI)	175 (128 – 222)	178 (52 – 303)
Total number of adult CF centres, *N*	18	10
Number of centres according to size		
<100 adults, *N* (%)	5 (27.8)	5 (50.0)
100–350 adults, *N* (%)	12 (66.7)	3 (30.0)
>350 adults, *N* (%)	1 (5.6)	2 (20.0)
Characteristics of adults:		
Age in years, median (IQR)	27 (22 – 35)	29 (23 – 37)
Female, *N* (%)	1422 (45.2)	803 (45.2)
Pancreatic insufficient,[†] *N* (%)	2582 (82.6) [a]	1427 (81.4) [b]
CF related diabetes, *N* (%)	1030 (32.7)	572 (32.2)
BMI in kg/m^2, median (IQR)	22.2 (20.2 – 24.8)	22.1 (19.8 – 24.5)
Annual review $\%FEV_1$,[‡] median (IQR)	67.6 (46.9 – 86.0)	64.4 (44.3 – 84.2)
Best $\%FEV_1$,[‡] median (IQR)	73.2 (53.5 – 91.0) [c]	69.2 (47.9 – 89.6) [d]

[§] These centres were included in the analyses to determine the centre-by-centre discrepancy between paired annual review and best $\%FEV_1$.

[Ω] These centres were excluded from the analyses to determine the centre-by-centre discrepancy between paired annual review and best $\%FEV_1$ due to the extent of missing best $\%FEV_1$ data.

[†] Data for pancreatic replacement therapy (PERT) use were obtained. People on PERT were considered 'pancreatic insufficient'. People not on PERT were considered 'pancreatic sufficient'. PERT use documented as 'unknown' is considered as missing data.

[‡] % predicted FEV_1 was calculated using the Knudson equation.

[a] Pancreatic status was missing for 23 (0.7%) of the adults in centres with ≥80% completeness for best FEV_1 data.

[b] Pancreatic status was missing for 24 (1.4%) of the adults in centres with <80% completeness for best FEV_1 data.

[c] Best FEV_1 data was available for 2780 (88.3%) of the adults in centres with ≥80% completeness for best FEV_1 data.

[d] Best FEV_1 data was available for 427 (24.0%) of the adults in centres with <80% completeness for best FEV_1 data.

Figure 1: A funnel plot and a graph showing the age and pancreatic status adjusted annual review %FEV_1 according to CF centre size in 2014

Figure 1a: Funnel plot

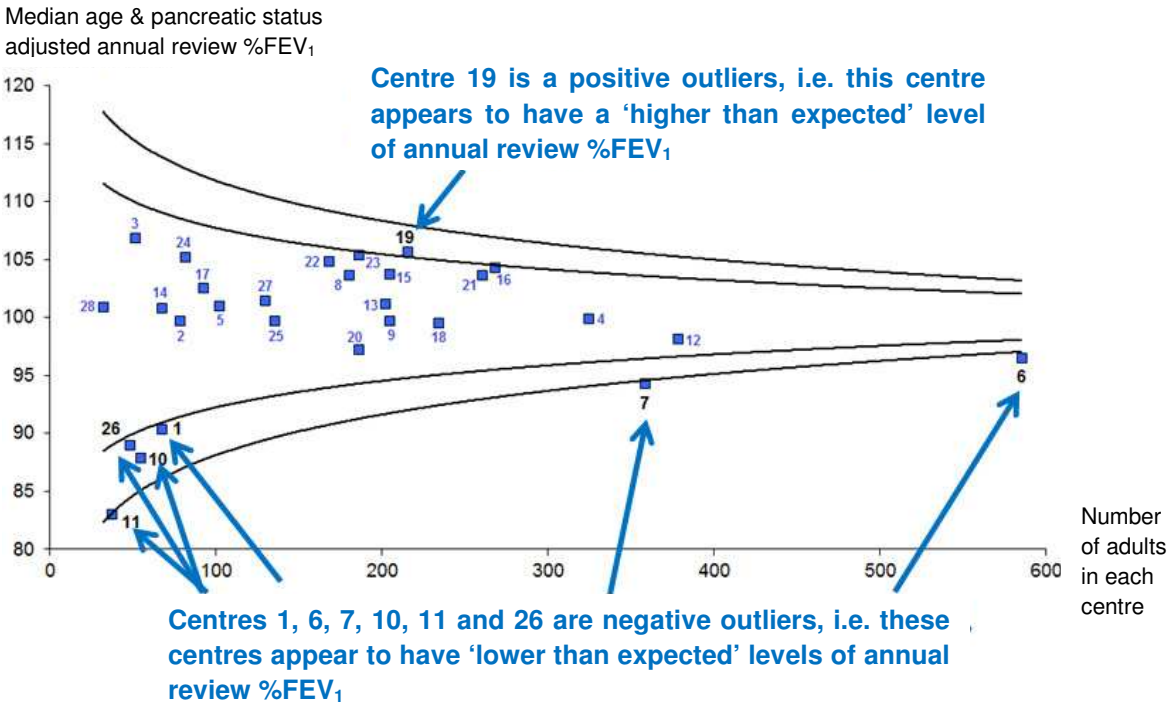

Figure 1b: Graph with a LOESS curve

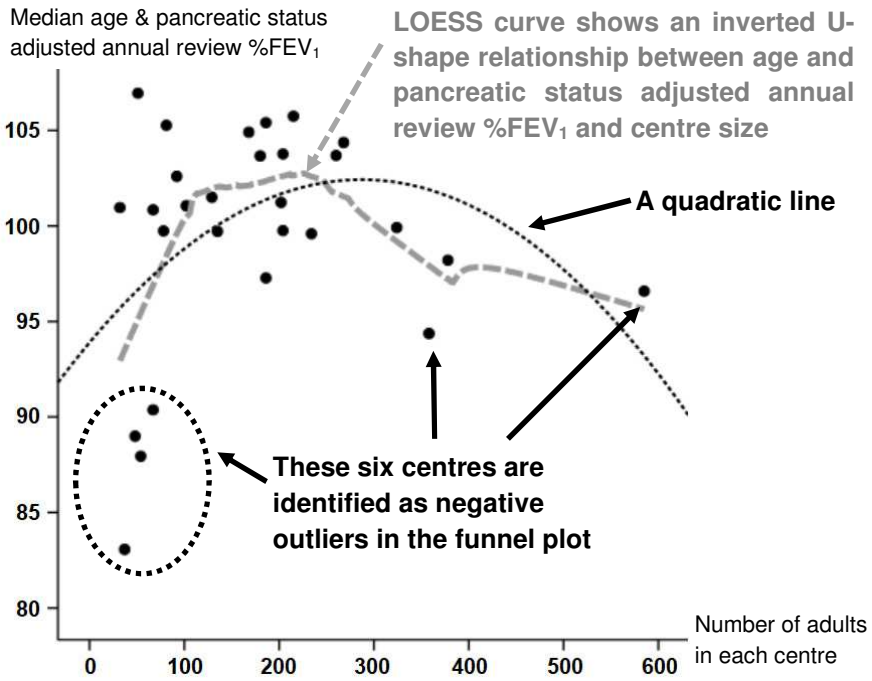

Figure 2: Graph showing the centre-level paired mean difference between annual review and best %FEV_1 according to the size of CF centres in 2014 (only data from centres with ≥80% data completeness for best FEV_1 included)