**Assessing the validity and applicability of the French 3-year prognostic score in the UK cystic fibrosis population**

**Authors:**

Dr. Freddy Frost (BMBS, BMedSci)

Adult CF Unit,

Liverpool Heart and Chest Hospital NHS Trust

Thomas Drive, Liverpool L14 3PE, United Kingdom

Email: Freddy.Frost@lhch.nhs.uk

Dr. Dilip Nazareth (MBBS, MSc, M.D., FRCP)

Adult CF Unit,

Liverpool Heart and Chest Hospital NHS Trust

Thomas Drive, Liverpool L14 3PE, United Kingdom

Email: Dilip.Nazareth@lhch.nhs.uk

Dr. Matthew Shaw (BSc, PhD)

Liverpool Heart and Chest Hospital NHS Trust

Thomas Drive, Liverpool, L14 3PE, UK

Email : Matthew.Shaw@lhch.nhs.uk

Dr. Mohamed Al-Aloul (BMSc (Hons), MBChB (Hons), MD, FRCP)

Consultant Respiratory & Transplant Physician

Manchester University NHS Foundation Trust

Southmoor Road, Wythenshawe, M23 9LT, United Kingdom

Email: Mohamed.Alaloul@mft.nhs.uk

**Authorship**: **M. Al-Aloul:** conceptualisation. **F. Frost, D. Nazareth and M. Al-Aloul** contributed equally to: Methodology, Data curation, Analysis, Writing- Original draft preparation, Writing-Reviewing and Editing. **M. Shaw**: Data curation, analysis and final draft preparation.

**Funding Source**: None

**Corresponding Author**: Dr Mohamed Al-Aloul

**Running title**: The validity of the French 3-year prognostic score in the UK CF population

**Keywords**: Cystic Fibrosis; Lung transplantation; Prognostic scores

**Abbreviations**: Lung transplantation (LT); Cystic fibrosis (CF); Forced expired volume in one second (FEV1); UK CF Registry (UK CFR); Burkholderia cepacia complex (BCC).

**Conflict-of-Interest**: None relevant

**Acknowledgements**: We acknowledge and thank the UK CF Trust for making the data available to us for analysis

**Funding**: None received

**Declaration of interest**: None

**Abstract:**

**Introduction**: Models that predict outcomes, aid prognostication and inform the assessment of urgency for lung transplantation (LT) in CF are in demand. A prognostic score derived from the French adult CF registry to predict death or LT over 3-year follow-up was described in 2017 and validated using Canadian CF registry data. We assessed its performance in the UK CF population.

**Methods**: The French prognostic score was applied to untransplanted adults with CF. The index year (2014) and outcomes (Death or LT) were evaluated to 2017. Receiver operator characteristics plots and area under curve (AUC) was computed.

**Results**: 4407 adults with CF met the inclusion criteria. After 3 years, 7.1% (p<0.001) were dead or had received LT compared to the French (12.8%) and Canadian (9.4%) cohorts. The French score deemed 592 (26.2%) “High Risk” - death/LT occurred in 189/592 (30.2%), less than previously reported in France and Canada (p<0.0001). The discriminatory power of the French score was lower (AUC 0.830) than reported. Recalibration yielded only marginal improvement in model performance (AUC 0.833).

**Discussion**: The French prognostic score does not perform as well in the UK as reported elsewhere. Bespoke UK scores are needed to aid prognostication and inform LT decision making.

**Introduction:**

Lung transplantation (LT) can offer improved prognosis and quality of life in carefully selected individuals with CF. In the UK, there is a mismatch between transplant demand and availability which can result in up to 30% of people with CF (pwCF) dying whilst awaiting transplantation [1]. There has therefore been much interest globally in developing models which can predict outcomes and aid prognostication thereby informing the assessment of urgency of need for transplant.

Several attempts at developing prognostic scores have been performed in pwCF but it has proven difficult to develop a metric that better predicts death than a forced expired volume in one second (FEV1) <30% predicted [2, 3] . Further, prognosis has dramatically improved over the past decades due to advances in integrated care provided by multidisciplinary teams in CF centres and latterly the availability of CFTR modulators; prognostic factors have thus changed over time and studies performed using data in previous eras may not be appropriate for current evaluation of pwCF.

A prognostic score derived from the French CF registry data to predict death or LT over a 3-year period in the adult CF population [Supplement E1] was described in 2017 [4] and subsequently validated in a North American setting using the Canadian CF registry [5]. This score is simple to calculate in clinic or at the bedside and hence there has been much interest in its use internationally.

Differences in healthcare systems, access to treatments and population genotype composition could all play a role in shaping patient prognosis. There was a striking resemblance in patient characteristics between the French and Canadian registries and both countries share comparable healthcare systems with similar access to transplantation. Before this score can be applied to wider international CF populations, it is important to validate it in a different healthcare setting. Our objective was therefore to comprehensively evaluate the French 3-year prognostic survival score using the UK CF Registry (UK CFR).

**Patients and Methods**:

**Data source:**

The UK CFR is a Research Ethics Committee approved (Huntingdon Research Ethics Committee 07/Q0104/2) database providing a longitudinal dataset of the UK CF population with excellent coverage and robust data quality assurance mechanisms [6]. The dataset includes longitudinal demographics, clinical characteristics, treatment information and outcomes including death and transplantation. All patients included in the UK CFR have consented to the collection and use of their medical information for research purposes. This study and subsequent data release were approved by the CF Registry Research Committee (CF Trust, UK).

**Patient selection:**

The index year from which baseline data were extracted to compute the individual risk scores was from January 1, to December 31, 2014. The follow up period was from January 1, 2015 through to December 31, 2017 inclusive. Patients were included for analysis when the following criteria were met: age ≥18 years, alive at December 31st, 2014, an annual review record in 2014 and were not lost to follow-up by the end of 2017. PwCF who received a LT prior to 2014 were excluded.

**Model variables:**

For each individual included in the study, the prognostic score put forward by Nkam et al [4], herein referred to as the “French Score”, see Table S1, was calculated and a risk category assigned as previously described. All variables were included from the index year only in keeping with prior derivation and validation studies. Annualised variables, representing the CF annual review year, e.g. intravenous antibiotic days, therefore represent the 12 months leading up to December 31, 2014. Death or transplant, the endpoints of interest, were recorded up until December 31st, 2017.Lung function was calculated as percent predicted forced expiratory volume in 1 second (ppFEV1) using Global Lung Initiative reference ranges [7]. Body mass index (BMI) was calculated as kg/m2.

**Model validation and statistical methods:**

The model was evaluated with measures for discrimination and calibration. The primary outcome of interest was the discriminatory ability of the French score for death or transplant within 3 years, measured as area under curve (AUC) of the receiver operator characteristic (ROC) plot. To assess for differences in study populations, a comparison of the model variable characteristics in each of these 3 populations was undertaken. An updated multivariable logistic regression model was constructed re-evaluating associations of the individual variables of the French score with death / LT in the UK setting. Model performance was assessed with the bootstrap method, where the complete datasets were sampled from repeatedly, and the final multivariate logistic regression model was refit 100 times. Model performance summary statistics were calculated for each iteration with the average across all the bootstrapped samples then calculated [8]. Calibration was assessed by plotting predicted risk against observed risk for each of the models’ risk strata, with intercept and slope calculated by linear regression. Miscalibration of the original model was adjusted for with the newly derived regression coefficients for each variable. Performance of the original and recalibrated models is described comparing risk predictions with observed outcomes for each risk category. All analyses were performed in RStudio (v1.0.1.136, RStudio Inc). Unadjusted p-values are presented throughout with statistical significance considered at <0.05.

**Results:**

Data were available for 9433 patients in the index year of 2014. Based on the original French study criteria, 5026 patients were excluded leaving 4407 included for analysis (Figure 1).

Baseline clinical characteristics of subjects included in the study are presented in Table 1. Briefly, median [IQR] age was 28.1 years [22.7, 35.7], 55.5% were male, ppFEV1 was 74.0% [54.0, 91.0] and BMI was 22.2 kg/m2 [20.2, 24.7]. After three years, 315/4407 (7.1%) were dead (n=217) or had received a lung transplant (n=98), which is significantly less than reported in the French and Canadian cohorts (12.8% and 9.4% respectively). As expected, those who died or received a lung transplant had lower lung function (39.0% [32.0 to 51.0] vs. 75.0 [58.0 to 92.0]), lower BMI (20.5 [18.5 to 22.8] vs. 22.3 [20.3 to 24.7] kg/m2) and received more intravenous antibiotics (0 [0 to 12] vs. 17 [0 to 52]); full comparison of clinical characteristics are presented in Table 1.

The French prognostic score (See online supplement E1) was calculated for all 4407 individuals. 2656/4407 (60.7%) were deemed “Low risk”, 1159/4407 (26.2%) “Moderate risk” and 592/4407 (13.4%) “High risk”. Comparative proportions in the respective risk categories in the French and Canadian populations were 51.4% vs 63.8%, 31.8% vs 23.5% and 16.8% vs 7.5%. Death or lung transplantation occurred in 2.0% of the “Low risk” cohort, 6.9% “Moderate risk”, and 36.6% of the “High risk” cohort. “Low Risk” and “Moderate Risk” outcomes were similar to those in the French and Canadian studies, but the UK “High Risk” cohort were less likely to achieve the primary outcome (54.5% and 55.0% in Canada and France respectively).

The ROC curve for the French prognostic score is presented in Figure 2. The AUC, equivalent to the c-statistic, of 0.830 was less than seen in the French (0.910) and Canadian (0.904) populations. When death and transplant were considered separately rather than as a composite outcome, the AUCs were 0.772 and 0.894 respectively, seeFigure 2. To assess overall model calibration, predicted risk was plotted against the observed risk, see Figure 3, with estimation of intercepts and slope by linear regression. Overall, the French score appeared to overestimate risk of death/transplantation in the UK cohort, particularly in the “high risk” strata.

Next, the individual components of the French prognostic score were included in a multivariable logistic regression model to assess their association with the primary outcome in our own CF cohort and are presented in Table 2. In general, variables had similar associations to those seen in the score’s derivation, with the exception of hospitalisation and number of exacerbations. The coefficients from this new model were then used to revise the scoring structure of the French model, see Table 2. Despite adjustment, only modest improvement in calibration slope from 0.65 to 0.69 was observed, with no improvement in model discrimination (AUC [95% CI]) from 0.830 [0.828 to 0.833] to 0.834 [0.831 to 0.836] for the recalibrated score.

Finally, to evaluate whether population differences explained the different results in the UK, model variable characteristics were compared between the three countries and are presented in Table 3. The median [IQR] ppFEV1 was higher in the UK population as compared to France (74% [54.0-91.0] and 58.3% [39.4–79.8] respectively) but not Canada (70.2% [51.5–88.0]). Similarly, differences were also observed in *Burkholderia cepacia complex* (BCC) prevalence between the UK and Canada (5.1% and 15.9% respectively) but not the UK and France (5.1% and 3.4% respectively). Other model variables were broadly similar across groups.

**Discussion**:

This study aimed to comprehensively evaluate the French prognostic score put forward by Nkam et al [4], in the UK cystic fibrosis setting. We found the French score was able to successfully identify pwCF at increased risk of death or transplantation. However, overall, the discriminatory power of the French score was less than previously reported in other CF populations.

The French score was recently validated in the Canadian CF population generating interest in its potential use internationally. However, the reduced AUC seen here (0.83 compared to ~0.91 seen in France and Canada), and inferior calibration suggest the score provides less reliable estimates of risk and may not be appropriate for ubiquitous application in the UK. Although an AUC of ~0.80 may be considered “excellent” in a clinical prediction tool [8], calibration of the original model was sub-optimal and if applied to the entire UK population, almost 1000 more people with CF would be mis-classified as high-risk than if a score with AUC ~0.90 was used. To further understand the discrepancy in score performance, we conducted a between-country comparison of the rates of death/transplantation amongst the three prognostic risk categories. This analysis revealed that the reduced prognostic performance appeared to be driven predominantly by the reduced proportion of UK “High Risk” individuals who met the death/transplantation composite outcome within 3 years.

In their Canadian validation paper, Coriati et al [5] pointed out the striking similarities in population characteristics and healthcare settings between their cohort and the French CF population. To investigate if differences in population characteristics could explain the differences in score performance in the UK we compared model variable characteristics across the three studies. All three study cohorts were broadly similar in terms of age, sex and BMI but interestingly, there were occasional differences between individual countries, for example lung function and BCC prevalence. The French cohort did include a higher prevalence of home oxygen and NIV use, potentially explained by either differing healthcare practices or a more co-morbid cohort in that study. Nevertheless, no consistent differences between the UK and both France and Canada existed, thus differences in populations alone cannot explain the differing model performance.

Whilst there are many similarities in the national health services of the UK, Canada and France, and all three countries follow international guidelines for the selection of lung transplant candidates, a stark difference exists in terms of access to transplant. For example, the UK (2.9 double Lung transplants per million population) falls well behind France (4.9 per million) and Canada (8.9 per million) [9]; as a consequence only 3-4% of adults with CF in the UK have received a lung transplant, compared to 11% in France [10]. These differences are reflected in the relative contribution of lung transplantation to the composite death/transplantation outcome, which was considerably lower in the UK data here (31.1%) than in the French data (79.5%). This disparity may contribute to differing score performance in the UK.

In support of the argument that the disparity of UK pwCF access to transplant is a factor attenuating score performance, we found the French score performance improved when considering transplant as an individual outcome of interest. Our interpretation of this finding is that the differences between countries in terms of the opportunity of a transplant mean that in the UK, only a super-selected cohort undergo transplantation. Consequently, nearly all of those undergoing transplant in the UK would score highly on the prognostic risk scale and would also have been likely to receive a transplant in the Canadian and French setting. This in turn presents apparent concordance between the model prediction and the actual “transplant” outcome. However, a considerable percentage of the high-risk UK CF population do not undergo lung transplantation where their French and Canadian counterparts would have done so. This shifts more onus onto the “Death” component of the outcome in the UK, but given the score only predicts survival to 3 years, any high scoring UK subject, who did not undergo transplant but survived longer than 3 years will reduce the predictive power of the score. Further work in datasets with longer follow-up is needed to determine whether 'high-risk' patients in the UK who do not undergo transplant within 3-years, go on to get transplanted in subsequent years.

An alternative explanation could be that variables derived and validated in the French model are not significantly correlated with death and/or transplantation in the UK CF population. However, including the French score variables in a multivariable logistic regression model revealed nearly all variables were relevant to these outcomes in our CF cohort. Some variables e.g. number of exacerbations and hospitalisations were not independent risk factors for death/transplantation in the UK. This may be due to variations in the definition of an exacerbation and local clinicians’ thresholds for inpatient therapy in different countries. Interestingly, when the composite death/transplantation outcome was split in the original French score derivation, these variables were not relevant to death and were instead predictors of transplantation, again supporting the notion that differences in transplant practices may explain some of the differences in score performance in the UK. Importantly, recalibration of the French score based on the UK regression model coefficients improved the reliability of the predictive score in the low and medium risk groups, but the model, continued to overestimate risk in the higher risk UK CF population and discrimination did not improve.

Overall, these findings highlight the complexities in comparing international datasets and suggest a UK specific score such as those recently put forward by Keogh et al [11] and Alaa et al [3] may be more relevant and further work is needed in this field.

It is important for clinicians and patients to understand prognosis with and without transplant such that an informed decision regarding the need to proceed to transplantation can be made. A test that could provide prognostic information for outcomes with and without transplant would be attractive. Indeed, the lung allocation score developed in the USA (US-LAS) in 2005 [12] and subsequently adopted by EuroTransplant in 2009 incorporates two sets of variables to produce a numeric score used to rank candidates for urgency of donor organ allocation: one set to predict waiting list mortality and another predicting 1 year post-operative survival. The UK lung transplant community has not adopted the US-LAS because this is weighted twice as heavily towards reducing waiting list mortality as it is for ensuring acceptable 1-year post transplant survival. NHS Blood and Transplant (NHSBT) Cardiothoracic Advisory Group Lung Allocation Working Party concluded that 1 year outcomes are too conservative and were eager to avoid the situation where diverting organs to the sickest may be at the expense of jeopardising good 3-5 year post transplant survival [Urgent Lung Allocation Scheme, in Cardiothoracic Advisory Group Meeting Minutes [*https://www.odt.nhs.uk/transplantation/cardiothoracic/cardiothoracic-advisory-group/*](https://www.odt.nhs.uk/transplantation/cardiothoracic/cardiothoracic-advisory-group/)2014, NHS Blood and Transplant: London].

Given the current median post-transplant survival for CF recipients is over 10 years [13], there is clearly a need to develop disease specific metrics balancing the need to reduce waiting list mortality (by timely listing and transplantation of the sickest) versus sustaining the good post-transplant outcomes observed to date [14]. To comprehensively assess pre and post-transplant outcomes it is likely UK CF Registry alone is insufficient and more specific post-transplant variables would be needed. Data linkage of CF and transplant registries would better facilitate this and may lead on to the development of a bespoke prognostic model that achieves the desired balanced outcomes mentioned above.

There are several limitations to consider for this study. We were unable to match some variables identically to the French score, for example, the UK CF Registry presents lung function using GLI reference equations [7] rather than Knudson equation [15] as reported by French and Canadian studies, however the Canadian study found no difference in score performance between Knudson and GLI derived predicted lung function so our results are unlikely to be confounded in that regard. Similarly Nkam et al [4], defined oral corticosteroid use as >3 months, however the UK CF Registry only captures the use of any corticosteroids in the previous 12 months. Reassuringly the Canadian registry captures data in a similar way to the UK and given the similarity in outcomes between French and Canadian groups, this limitation is unlikely to explain the differences seen here. Those with incomplete follow up data were excluded from our analysis, in keeping with the methods used in the derivation of the French score. This represents a potential source of bias: it is possible that these pwCF are sicker or non-adherent and therefore less likely to be considered viable transplant candidates, hence for those pwCF only the second outcome of death is relevant. In other words, that group may be at different risk compared to those with complete follow up data. Reassuringly the proportion excluded here (8.5%) is similar to that in the original derivation cohort making gross systematic differences unlikely.

The strengths of this study lie in the replication of methods used for cohort derivation in both prior studies and a high level of data fidelity - regular quality audits via a data validation programme, have confirmed a high level of data accuracy and completeness of UKCFR data (98.8%) [16]. Moreover, this is the largest study to date of the French prognostic score.

To conclude, the French CF prognostic score put forward by Nkam et al [4] does not perform as well in the UK setting as previously reported in France and Canada. This finding is most likely as a result of poorer access to lung transplantation in the UK. Bespoke prognostic scores incorporating pre and post-transplant survival are needed to help guide clinicians, patients and transplant governing bodies in the UK.

**References**

1. Kourliouros, A., R. Hogg, J. Mehew, et al., *Patient outcomes from time of listing for lung transplantation in the UK: are there disease-specific differences?* Thorax, 2019. **74**(1): p. 60-68.

2. Ramos, K.J., P.J. Smith, E.F. McKone, et al., *Lung transplant referral for individuals with cystic fibrosis: Cystic Fibrosis Foundation consensus guidelines.* J Cyst Fibros, 2019. **18**(3): p. 321-333.

3. Alaa, A.M. and M. van der Schaar, *Prognostication and Risk Factors for Cystic Fibrosis via Automated Machine Learning.* Sci Rep, 2018. **8**(1): p. 11242.

4. Nkam, L., J. Lambert, A. Latouche, G. Bellis, P.R. Burgel, and M.N. Hocine, *A 3-year prognostic score for adults with cystic fibrosis.* J Cyst Fibros, 2017. **16**(6): p. 702-708.

5. Coriati, A., J. Sykes, L. Nkam, M.N. Hocine, P.R. Burgel, and A.L. Stephenson, *Validation of the French 3-year prognostic score using the Canadian Cystic Fibrosis registry.* J Cyst Fibros, 2019. **18**(3): p. 396-398.

6. Taylor-Robinson, D., O. Archangelidi, S.B. Carr, et al., *Data Resource Profile: The UK Cystic Fibrosis Registry.* Int J Epidemiol, 2018. **47**(1): p. 9-10e.

7. Quanjer, P.H., S. Stanojevic, T.J. Cole, et al., *Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations.* Eur Respir J, 2012. **40**(6): p. 1324-43.

8. Hosmer, D.W. and S. Lemeshow, *Applied Logistic Regression, Second Edition*. 2000: John Wiley & Sons, Inc.

9. Matesanz, R., *International Figures on Donation and Transplanttaion - 2013*, in *NEWSLETTER TRANSPLANT*. 2014: Madrid.

10. Zolin A, O.A., Naehrlich L, van Rens *ECFSPR Annual Report 2017*. 2019: Denmark.

11. Keogh, R.H., S.R. Seaman, J.K. Barrett, D. Taylor-Robinson, and R. Szczesniak, *Dynamic Prediction of Survival in Cystic Fibrosis: A Landmarking Analysis Using UK Patient Registry Data.* Epidemiology, 2019. **30**(1): p. 29-37.

12. Egan, T.M., S. Murray, R.T. Bustami, et al., *Development of the new lung allocation system in the United States.* Am J Transplant, 2006. **6**(5 Pt 2): p. 1212-27.

13. Stephenson, A.L., J. Sykes, Y. Berthiaume, et al., *Clinical and demographic factors associated with post-lung transplantation survival in individuals with cystic fibrosis.* J Heart Lung Transplant, 2015. **34**(9): p. 1139-45.

14. Koutsokera, A., R.A. Varughese, J. Sykes, et al., *Pre-transplant factors associated with mortality after lung transplantation in cystic fibrosis: A systematic review and meta-analysis.* J Cyst Fibros, 2019. **18**(3): p. 407-415.

15. Knudson, R.J., M.D. Lebowitz, C.J. Holberg, and B. Burrows, *Changes in the normal maximal expiratory flow-volume curve with growth and aging.* Am Rev Respir Dis, 1983. **127**(6): p. 725-34.

16. Gunn, E., N. Caine, R. Cosgriff, S.C. Charman, and S.B. Carr, *P082 UK Cystic Fibrosis Registry data validation programme.* Journal of Cystic Fibrosis, 2019. **18**: p. S80-S81.