


RESEARCH ARTICLE

Open Access



Prospective multicenter randomized patient recruitment and sample collection to enable future measurements of sputum biomarkers of inflammation in an observational study of cystic fibrosis

Theodore G. Liou^{1,2*} , Frederick R. Adler^{3,4}, Natalia Argel⁵, Fadi Asfour², Perry S. Brown⁶, Barbara A. Chatfield², Cori L. Daines⁷, Dixie Durham⁶, Jessica A. Francis¹, Barbara Glover⁸, Theresa Heynekamp⁹, John R. Hoidal¹, Judy L. Jensen¹, Ruth Keogh¹⁰, Carol M. Kopecky¹¹, Noah Lechtzin¹², Yanping Li¹, Jerimiah Lysinger¹³, Osmara Molina⁷, Craig Nakamura⁸, Kristyn A. Packer¹, Katie R. Poch¹⁴, Alexandra L. Quittner^{15,16}, Peggy Radford⁵, Abby J. Redway⁹, Scott D. Sagel¹¹, Shawna Sprandel¹³, Jennifer L. Taylor-Cousar^{14,17}, Jane B. Vroom^{1,2}, Ryan Yoshikawa⁸, John P. Clancy¹⁸, J. Stuart Elborn¹⁹, Kenneth N. Olivier²⁰ and David R. Cox²¹

Abstract

Background: Biomarkers of inflammation predictive of cystic fibrosis (CF) disease outcomes would increase the power of clinical trials and contribute to better personalization of clinical assessments. A representative patient cohort would improve searching for believable, generalizable, reproducible and accurate biomarkers.

Methods: We recruited patients from Mountain West CF Consortium (MWCFC) care centers for prospective observational study of sputum biomarkers of inflammation. After informed consent, centers enrolled randomly selected patients with CF who were clinically stable sputum producers, 12 years of age and older, without previous organ transplantation.

(Continued on next page)

* Correspondence: ted.liou@utah.edu

In Memoriam of Barbara A Chatfield: August 25, 1959—October 8, 2016.
Authors are alphabetical except first (TGL), advisory committee (JPC, JSE, KNO) and senior (DRC).

¹Adult Cystic Fibrosis Center, Division of Respiratory, Critical Care and Occupational Pulmonary Medicine, Department of Internal Medicine, University of Utah, 26 North Mario Capecchi Drive, Salt Lake City, UT 84132, USA

²Intermountain Pediatric Cystic Fibrosis Center, Division of Pediatric Pulmonology, Department of Pediatrics, University of Utah, 81 North Mario Capecchi Drive, Salt Lake City, UT 84113, USA

Full list of author information is available at the end of the article



(Continued from previous page)

Results: From December 8, 2014 through January 16, 2016, we enrolled 114 patients (53 male) with CF with continuing data collection. Baseline characteristics included mean age 27 years (SD = 12), 80% predicted forced expiratory volume in 1 s (SD = 23%), 1.0 prior year pulmonary exacerbations (SD = 1.2), home elevation 328 m (SD = 112) above sea level. Compared with other patients in the US CF Foundation Patient Registry (CFFPR) in 2014, MWCFC patients had similar distribution of sex, age, lung function, weight and rates of exacerbations, diabetes, pancreatic insufficiency, CF-related arthropathy and airway infections including methicillin-sensitive or -resistant *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Burkholderia cepacia* complex, fungal and non-tuberculous *Mycobacteria* infections. They received CF-specific treatments at similar frequencies.

Conclusions: Randomly-selected, sputum-producing patients within the MWCFC represent sputum-producing patients in the CFFPR. They have similar characteristics, lung function and frequencies of pulmonary exacerbations, microbial infections and use of CF-specific treatments. These findings will plausibly make future interpretations of quantitative measurements of inflammatory biomarkers generalizable to sputum-producing patients in the CFFPR.

Keywords: Cystic fibrosis, Randomized observational trial, Study design, Sputum inflammation, HMGB-1, Neutrophil elastase, Calprotectin, Cystic Fibrosis Foundation patient registry

Background

Neutrophil-predominant, intense airway inflammation associated with chronic airway infections drives CF disease. Multiple molecules [1–4] describe and mediate airway inflammation in CF. These biochemical markers of airway inflammation may mark disease progression and potentially provide surrogate markers of survival or intermediate outcomes that could improve the power of clinical trials to demonstrate efficacy of novel treatments [2, 3, 5–8]. Lung function, pulmonary exacerbations and weight are the best clinical predictors of survival [9, 10] and are the key end-points for multiple trials with high impact on CF treatment [11–17]. However, the insensitivity of these clinical markers requires large trial enrollments and prolonged observation to ascertain effects [11, 13, 15, 18]. Understanding relationships between these clinical features of disease and the laboratory measurements of biomarkers of inflammation will improve our knowledge of pathophysiology and enable better understanding of the impacts of current and investigational treatments.

Most evaluated inflammatory signals are increased in CF airways relative to normal airways [3, 19–21]. Three potentially useful biomarkers are associated with disease progression. High mobility group box-1 protein (HMGB-1), a highly conserved protein [22, 23], is associated with pulmonary exacerbations [24]. Sputum concentration correlates with concurrent lung function and number of prior-year exacerbations [25]. HMGB-1 sputum concentration predicts time to next exacerbation, number of subsequent pulmonary exacerbations and time to death or lung transplantation in CF. [25] Calprotectin, a proinflammatory molecule found abundantly in CF sputum, decreases with antibiotic treatment; high levels are associated with more rapid recurrence of a pulmonary exacerbation [26, 27]. Neutrophil elastase activity (NE) appears in airway secretions in CF [28],

induces interleukin-(IL)-8 production [29], and predicts accelerated lung function loss in children with CF [5]. Finally, granulocyte macrophage colony stimulating factor (GM-CSF) is another potentially useful biomarker. Its concentration in sputum at the beginning of hospitalization for a pulmonary exacerbation was strongly associated with the acute decrease in forced expiratory volume in 1 s (FEV₁) normalized to percent predicted FEV₁ for age, height, sex, race and ethnicity (FEV₁%) [25, 30]. GM-CSF may thus be an objective measurement of severity of an exacerbation.

However, all studies of these four potential biomarkers in CF were performed in single centers among patients recruited in ways subject to observer biases [5, 24, 25, 27]. Comparisons of predictive ability for future clinically relevant events between HMGB-1, calprotectin and neutrophil elastase cannot be done with existing data. We designed and implemented an observational study in the Mountain West CF Consortium (MWCFC) to allow direct comparison of the potential biomarkers and reduce biases. Because this study could facilitate the assessment of other clinically relevant conditions in CF and their relationship to inflammation with relatively small incremental costs, we included additional biomarkers and data collections in our study design to evaluate combinations of the primary molecules with each other and with other inflammatory markers to understand both our primary clinical event, pulmonary exacerbations, and additional secondary clinical conditions of high importance.

We report here the nature and implementation of our study design centered on randomized patient selection. This methodology strives to minimize observer bias in recruitment to maximize believability and generalizability of analysis [31, 32]. To demonstrate success, we compared cohort characteristics with patients in the CF Foundation

Patient Registry (CFFPR) [33] in 2014 who would have fulfilled inclusion and exclusion criteria.

Methods

Study setting

Nine CF Care Centers accredited by the US CF Foundation and located in the Mountain West Region participated in this trial. The Mountain West Region includes (from North to South and West to East) Idaho, Montana, Wyoming, Nevada, Utah, Colorado, Arizona, and New Mexico. MWCFC centers provide care to about 10% of the patients with CF [34] who live in approximately 25% of the total land area of all 50 of the United States [35].

Pre-study preparations

Prior to the general study planning meeting, the initial investigators (TGL, FRA, JLJ, RK, DRC) held multiple discussions to develop core statistical and logistical study plans. At the meeting, RK and DRC taught on core design issues (Additional files 1, 2 and 3, Video: Principles of Study Design: A Conversation Between David Cox and Ruth Keogh and Additional file 4: Transcript of Video) to improve the understanding of all investigators for study procedures. Two non-MWCFC collaborators provided guidance on additional important design elements (NL, ALQ), and three clinical investigators served as an external advisory committee (JPC, JSE, KNO) to review and modify procedures and provide oversight.

All personnel were trained at the University of Utah in study background, goals, inclusion and exclusion criteria, good clinical practice, patient safety and study procedures, data entry and records security. Research coordinators and laboratory personnel were trained individually at the University of Utah on sputum processing. The overall principal investigator (PI, TGL) and research manager (JLJ) performed on-site study initiation and training of investigators, study coordinators and laboratory technicians. Additional training was provided on-site or in Utah as

requested. Training records were archived at each site and at the University of Utah.

Study endpoints

The primary event for this observational trial is first pulmonary exacerbation following enrollment. The investigators at the general study planning meeting discussed previously published definitions for pulmonary exacerbations. Many clinical trials include a decision to start parenteral antibiotics in the criteria in order to facilitate retrospective identification of exacerbations [11]. After careful consideration, we decided to use the definition used in our original single center study (Table 1) [25] for several reasons: (A) we intended our primary event to be determined in a prospective manner, (B) a definition including use of parenteral antibiotics could have introduced bias in determination of a pulmonary exacerbation at the point of care, (C) the requirement for antibiotics might have been viewed as requiring an intervention despite ours being an observational study, and (D) a central goal was replication of results from the original single center study [25].

We considered using a scoring system to diagnose a pulmonary exacerbation [36] but felt that the additional complexity might reduce reporting of events especially for patients evaluated by clinicians not participating directly in the study. Because patients can be unstable but not quite reach the severity of disease described by our definition (Table 1), we defined a category of “mild exacerbation” to reduce the likelihood of arbitrary misclassification and to allow future investigation of this poorly defined clinical state.

The investigators and the external advisory committee reached unanimous consensus agreement with the definitions prior to enrollment of patients. These definitions were included in the application for investigational review board (IRB) or equivalent ethical oversight organization at all participating centers in the MWCFC.

Table 1 Symptoms and signs of a pulmonary exacerbation of CF^a

Symptoms	Signs
increased sputum, cough, dyspnea	10% drop in FEV ₁ or forced vital capacity
chest pain or tightness	Temperature > 38.4 °C
hemoptysis	Witnessed hemoptysis greater than 100 ml per episode
fever	SaO ₂ < 90% or PaO ₂ < 60 mmHg despite usual oxygen
chills	For adolescents a drop in SaO ₂ of 5% (for example, 97 to 92%)
arthralgias	Increased supplemental oxygen requirements
fatigue	Unplanned weight loss ≥5% of baseline body weight over 3 months
Other Considerations	
Respiratory arrest or failure requiring mechanical ventilation regardless of other criteria	
At site PI discretion for borderline cases or for cases with serious findings not included here	

^aA pulmonary exacerbation is defined as the presence of one symptom and one objective finding

Because there are multiple clinical questions for which insufficient information exists to allow independent studies, we discussed potential additional data collections and laboratory assessments. We incorporated additional clinical annotations and sample analyses that were judged feasible and that would have minimal potential to introduce bias or alter enrollment in our primary study.

Enrollment

After obtaining IRB approvals, we reviewed center-specific potential participant lists. Each center identified patients with confirmed CF either (A) older than 18 years of age able to provide informed consent and to expectorate sputum or (B) older than 12 (but less than 18) years able to provide assent and able to expectorate sputum or successfully undergo sputum induction as part of normal care, with parents or guardians able to consent. We enrolled patients after written informed consent only when judged clinically stable by enrolling investigators. We excluded patients unable or unwilling to provide sputum or who had previously received organ transplantation, were on immunosuppression beyond oral prednisone or on treatment with immunologic-based biologic therapies. We excluded pregnant women, prisoners and other vulnerable patients from enrollment but did not require withdrawal if a patient became vulnerable because of the non-coercive nature of continued participation in an observational study. Patients, guardians or investigators could request withdrawal at any time.

Power calculations

We sampled prior data and bootstrapped models to estimate the number of patients required to detect associations for HMGB-1 concentrations, our primary biomarker for study, with pulmonary exacerbations and survival outcomes [25, 37]. We additionally calculated the number of patients needed to study associations between GMCSF concentration and acute drops in FEV₁% at the beginning of a pulmonary exacerbation [25] and NE with drops in

FEV₁% among children [5] over two years of follow up (Table 2). Based on these calculations, we planned a minimum enrollment of at least 40 patients to explore the ability of our primary biomarker, HMGB-1, to predict our primary event, time to first pulmonary exacerbation. However, we hoped to enroll as many as 125 to a maximum of 175 patients, including both adults and children, to enable a believable and financially feasible analysis (see Additional files 1, 2, 3 and 4) of GMCSF as a reporter of severity of a pulmonary exacerbation [25].

Randomized selection of patients

To minimize observer bias (Additional files 2 and 4) [31, 32] and enroll a representative sample of patients with CF, we assigned a randomly chosen letter of the alphabet to each potentially eligible patient using one adult and one pediatric block per MWCFC center. Based on a prior recruitment rate of 53% [25], and allowing for a 10% clinic no-show rate, we chose a letter of the alphabet to assign to each adult and pediatric center to serve as the threshold between random inclusion or exclusion from enrollment. We recruited eligible patients attending clinic who had been assigned personal random letters earlier in the alphabet than the threshold letter for the center. The method eliminated the need for research coordinators to visit clinic for patients with assigned letters later in the alphabet than the center's threshold letter and allowed pre-clinic notification of potential enrollments and sample collections to clinical and laboratory personnel.

To maintain proportional enrollment among centers and reduce seasonal biases, we monitored enrollment rates at each center and adjusted the threshold letter earlier in the alphabet for enrollment that was too rapid and later in the alphabet for enrollment that was too slow after the first month and quarterly thereafter beginning with study month three. This enrollment speed adjustment strategy maintained randomization.

Table 2 Bootstrapped Power Calculations^a

Row	Statistical Model	Outcome Variable	Biomarker, Concentration or Activity	Percent Power	α-Level	Estimate of Patients Required
1	Proportional Hazards	Time to first Pulmonary Exacerbation	HMGB-1	90	0.01	40
2	Proportional Hazards	Time to Lung Transplantation or Death	HMGB-1	90	0.01	30
3	Linear Regression	Acute FEV ₁ % Drop with onset of a Pulmonary Exacerbation	GMCSF	80	0.01	175
4	Linear Regression	Acute FEV ₁ % Drop with onset of a Pulmonary Exacerbation	GMCSF	85	0.05	125
5	Linear Regression	FEV ₁ % Drop over 2 Years	NE	80	0.05	32

^aBased on prior published results, we sampled patients and bootstrapped statistical models of predictor biomarkers for outcome variables, setting percent power and α-level in order to estimate the number of patients similar to prior patients needed to detect associations. See the study protocol for more details (Additional file 5)

We required clinical stability and a sputum sample at enrollment. We defined clinical stability as the absence of a pulmonary exacerbation (Table 1) [25]. Because biomarker values change with time following an exacerbation, we collected the dates of the five exacerbations (if any) prior to the study enrollment date to enable adjustment for duration of clinical stability prior to sample collection. Guidance and examples were presented at the general study planning meeting and in the study protocol (Additional file 5, Study Protocol). Site PIs adjudicated borderline cases.

Shipping controls

Prior to finalizing sputum handling protocols, we examined the effects of shipping prior to initial sample processing on HMGB-1 measurements. We recruited 10 adult patients and collected expectorated sputum in Utah on a single day after informed written consent. We divided samples in two and processed half immediately to isolate an aqueous fraction (see next section or Additional files 6 and 7 for specific details). We added protease inhibitor cocktail (Sigma, St. Louis, MO, USA) and froze at -70°C until ready to perform HMGB-1 assay. We kept the second half of each sample on ice and shipped to ourselves overnight at 4°C (Fedex, Memphis, TN, USA), isolated the aqueous fraction, added protease inhibitor and froze. We measured HMGB-1 levels in both immediately-processed and shipped samples by ELISA using commercially-available antibodies (R&D Systems, Inc., Minneapolis, MN) [25, 38]. Results (see below) prompted incorporation of on-site sputum processing at each center.

Sputum collection, processing and shipping

We collected sputum after expectoration. We encouraged collection by alternating between research and clinical sample vials to minimize the disturbance of collection for cultures, but some research samples were collected after clinical samples [39]. We allowed sputum induction where that method was part of usual pediatric care. In accordance with Therapeutic Development Network standard operating procedure [40], we allowed one hour to collect sputum, transport from collection point on ice and initiate laboratory processing with an additional 1–3 h to complete processing. We further specified that during the initial hour, collection time was limited to 20 min, and 40 min were allowed to transport specimens on ice and initiate laboratory processing.

Samples were collected in 50 ml conical tubes supplied by the central study coordinators at the University of Utah. Following collection, samples were placed on ice and transported to the local processing laboratory. Sputum samples were weighed, diluted 1:1 with Hanks Buffered Saline Solution (HBSS, Sigma, St. Louis, MO)

and vortex mixed for 1 min. A sterile disposable pipette was used to transfer 0.25 ml of the vortex mixed sample to a 1.8 ml tube containing Streck Solution (Streck Inc. La Vista, NE); contents were mixed using the pipette, and the tube was sent via Fedex (Memphis, TN) to the U of Utah for cell counts and differentials. The remainder of the vortex mixed sputum sample was centrifuged for 20 min at 2800 g at 4°C to produce a top lipid layer, a middle aqueous layer and a bottom pellet layer. The layers were carefully separated by transfer pipettes; all personnel were trained to separate the layers in a way to avoid contamination of the aqueous fraction by the other two layers.

The lipid layer was carefully transferred to a new 1.8 ml cryovial, labeled with the participant study identification number (ID), collection date, fraction identifier (L) and sample number. The aqueous layer was divided in two. The first half was diluted with HBSS (Sigma) 1:1, vortex mixed for 10 s, aliquoted and labeled with participant study ID, date, fraction identifier (SA), sample number and aliquot number. The second half was diluted with protease inhibitor cocktail (Sigma) 1:1, vortex mixed for 10 s, aliquoted and labeled with participant study ID, date, fraction identifier (SB), sample number and aliquot number. The pellet was transferred to a 1.8 ml cryovial and labeled with participant study ID, date, fraction identifier (P), sample number and aliquot number. (See Additional files 6 and 7 for illustrated instructions for sample processing.)

Fractionated samples were frozen and stored at -70°C until shipping. Frozen samples were shipped on dry ice in split batches to prevent loss of any single sample (see Additional file 6) and immediately stored at -70°C in Utah until laboratory analysis.

We collected multiple samples from participants to address secondary study endpoints. To understand changes in airway inflammation, we collected sputum within 48 h of diagnosis of the first pulmonary exacerbation following enrollment and a convalescent sputum within 4–12 weeks of exacerbation onset. We encouraged collection of additional pulmonary exacerbation samples, particularly if a sample was missed at the time of first pulmonary exacerbation following enrollment. A final sample was requested at the end of the study whether the patient was acutely ill or stable.

A laboratory flow sheet with illustrated processing instructions was completed for each sample to ensure standardized processing and enable monitoring of individual samples (Additional file 6, Sputum Processing Instructions and Form). Written and video sputum processing instructions were available throughout the study (Additional file 5 and Additional file 7, Video: Cystic Fibrosis Biomarker Study: Sputum Collection and Processing).

Clinical annotation

Research coordinators collected and recorded information simultaneously with sputum sample collections including demographics, CFFPR identification numbers, home address, number and dates of CF exacerbations in the year prior to enrollment, other lung disease, smoking, menstrual and pregnancy histories, treatments, physical exam findings, spirometry results, bacterial and alcohol and acid-fast culture results and questionnaire results including Borg dyspnea index [41], GAD-7 anxiety scale [42, 43], PHQ-8 depression scale [43, 44], brief pain index [45], selected Munich Chronotype Questionnaire [46] items (MCQ, written permission for use obtained from Dr. Roenneberg) and food insecurity [47], and dates of death or lung transplantation. Many of these data addressed exploratory endpoints deemed interesting and feasible and unlikely to bias enrollments by the investigators (Table 3). All data were stored using the Research Electronic Data Capture (REDCap) system [48].

We reviewed REDCap and CF Foundation Patient Registry data for enrollment years to identify inconsistencies in study data. Study coordinators from Utah performed interim and end-of-study monitoring on-site to compare entered data with electronic medical records for every participant. Discrepancies were resolved by site study personnel prior to analyses.

Statistical analysis

We use the R statistical system [49]. We used patient home addresses and *The National Map* [50] to derive longitude, latitude and elevation. We calculated GAD-7 and PHQ-8 total scores [42, 44], weight-for-age z-score [9], FEV₁% [30, 51], and 5-year predicted survival [9]. We used χ^2 , *t*-test, linear regression and quasi-Poisson regression as appropriate to compare MWCFC with CFFPR data adjusted for study site, age, sex and other patient characteristics. We tested analyses for sensitivity to CF-specific treatments.

Table 3 Data Collections for Exploratory Endpoints

Endpoints potentially related to inflammation	Pertinent Clinical Data Collected
Depression, Anxiety, Pain	PHQ-8 Depression Scale GAD-7 Anxiety scale Borg Dyspnea Index Brief Pain Index
Environmental Factors: Altitude, Air Pollution, Climate	Sputum Collection Date Home Address (Latitude, Longitude, Altitude) Microbiological Sputum Culture Results
Infection Status	Sputum Culture results
Sleep and Circadian Rhythm: Sleep Phase and Duration	Munich Chronotype Questionnaire
Menstrual Cycle	Last Menstrual Period Date
Food Insecurity	History of missed or potentially missed meals GAD-7 anxiety scale Food Availability focused questions

Results

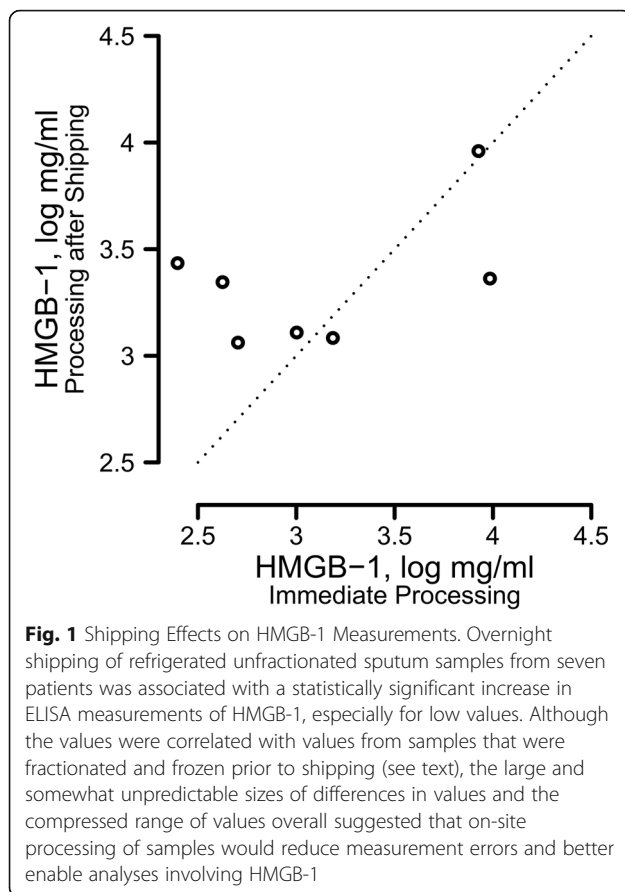
Shipping controls

Among 10 volunteers, 7 produced sufficient sputum for evaluation. After overnight shipping at 4 °C, HMGB-1 levels were elevated (linear regression: coefficient = 1.15, intercept = 5.19, *p* = 0.001) by factors as high as 2.8 (mean elevation factor 1.40, 95% CI = 0–2.8) compared to paired levels from sputum processed the same day as collection (Fig. 1).

Patients

Investigators enrolled 116 of the 154 children and 696 adults with confirmed CF in the MWCFC able to expectorate sputum; two patients were too young and were excluded after enrollment (Table 4). First enrollment was December 8, 2014; last enrollment was January 16, 2016 (Fig. 2). Five patients did not complete the sample collection portions of the study due to lung transplantation, death or other cause of loss of contact between study center and the patient. However, all 114 patients allowed long term follow up by providing permission to use their specific CF Foundation Patient Registry ID numbers. Queries specific to the study are ongoing for end-of-study data (Table 5).

Enrollment ranged by study site from 0 to 43% and 1 to 28% of eligible children and adults, respectively. The proportion of enrollment was unrelated to center size for pediatric, adult or combined age groups (linear regression, *p* > 0.1, all groups). Using linear, logistic or quasi-Poisson regression as appropriate for the outcome variable assessed, we evaluated seasonal effects on enrollment. Relative to Summer enrollment, patients enrolled in the Fall were 8.7 years older (*p* = 0.02) while patients enrolled during the Winter had an FEV₁% 15.9 percentage points higher (*p* = 0.004). There were no significant relationships for sex, weight-for-age or height-for-age [52] z-scores, number of prior year pulmonary exacerbations, diabetes, pancreatic sufficiency,



5-year predicted survival [9], anxiety or depression scores, any infection status or patient home altitude. Despite statistically significant findings for characteristics considered individually, no seasonal differences remained significant after Bonferroni correction.

Comparison of patient characteristics with National Data

The 2014 CFFPR reports 14,394 patients able to expectorate sputum, 12 years of age or older. We found no significant differences in disease characteristics, microbiology results or prescribed treatments compared with MWCFC patients (Table 4).

Sputum samples

Two shipments of samples (out of approximately 30 total shipments to date for the study) arrived in Utah partially thawed despite verified adherence to shipping procedures and appropriate packing with dry ice at the two originating centers. However, because we required portions of every sample to be split at the time of on-site processing as safeguards against shipping problems, unthawed aliquots of every sample collected were available for laboratory analyses.

The natural log of the average cell count per milliliter was 16.2 (SE 8.11) based on 112 of 114 enrollment

samples of sufficient volume without oral contamination defined as ≥ 5 squamous epithelial cells per low power light microscopy field on modified Wright's stain (see Additional file 5). The average differential of 36% neutrophils, 30% lymphocytes, $< 1\%$ eosinophils and 34% other cells was based on 108 out of 112 samples. A few samples had inadequate preservation in transport or sparse cellularity thus less than 200 cells were counted, but we counted 200–500 cells for the rest of the samples in accordance with Therapeutic Development Network standard operating procedures for samples from patients with CF collected for inflammatory marker measurements (mean 250 cells/sample, range 41–541) [40]. Total processing time from initiation of sputum collection to completion of processing was a mean of 54 min (SD = 1 h 38 min, max time = 3 h 45 min).

A total of 267 sputum samples were collected from patients. There were 116 enrollment samples, but two of those patients were excluded because they were too young to participate. The primary outcome for the study is time-to-first pulmonary exacerbation, and we captured the time-to-event data for the 92 patients who had these events during follow up.

To address secondary outcomes, we successfully collected additional data at the time of the first exacerbation during study follow up for 92 of the 114 enrolled patients and collected sputum samples for 52 of the events. Two additional sputum samples were collected in patients with subsequent pulmonary exacerbations where the first exacerbation sample was missed. The major reasons for missing first exacerbation sputum samples were diagnosis away from the CF Care Center and lack of ability to produce sputum within the 48 h collection window.

Discussion

We implemented a design that focused on randomized patient selection for a multicenter observational study of inflammatory sputum biomarkers in CF. We achieved our primary enrollment goal of recruiting a cohort representative of patients in the US with confirmed CF, 12 years of age or older and able to produce sputum. Age, sex, lung function, growth and nutrition factors, major morbidities including pancreatic insufficiency, CF related diabetes and airway infection prevalence and treatment frequencies were similar between MWCFC and national patients (Table 4). We achieved our secondary goal of minimizing seasonal and care-center size effects on patient selection. Randomization reduces observer-dependent selection bias and increases believability of data collection, generalizability of results and security of interpretation (Additional files 1, 2, 3 and 4) [31, 32].

Our investigation of the effect of shipping on unprocessed samples found a potential source of inaccuracy

Table 4 Comparison of Annualized CFFPR Data between MWCFC and other US Patients

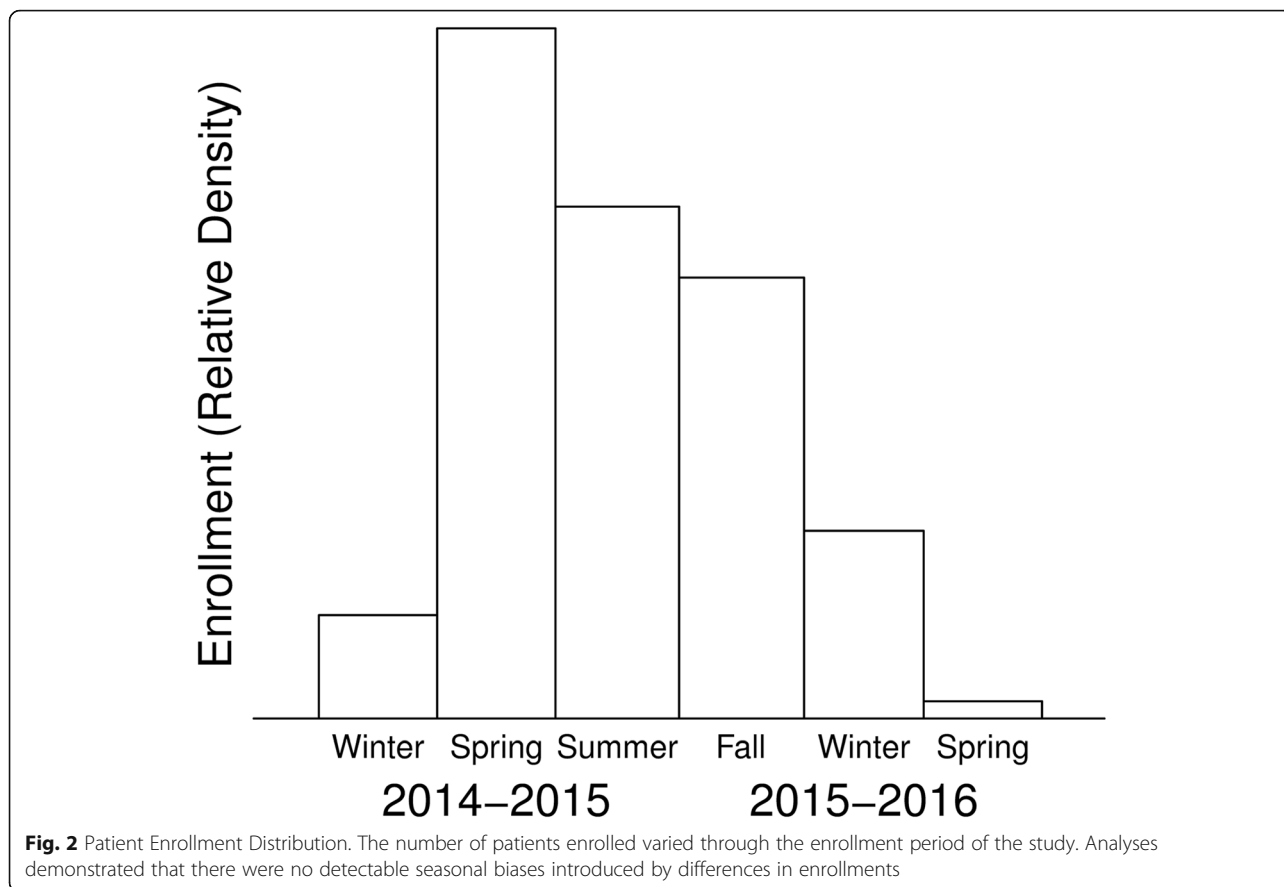
Patient Characteristics	MWCFC, <i>n</i> = 114		CFFPR 2014 <i>n</i> = 14,394	<i>p</i>
	Enrollment	Annualized for 2014		
Male sex, fraction	0.46	0.46	0.48	0.79
Age, Years, mean (SD)	28 (12)	27 (12)	27 (12)	0.71
FEV ₁ , Percent Predicted, mean (SD)	70 (22)	80 (23)	77 (27)	0.094
Height, cm, mean (SD)	167 (9.95)	167 (10.2)	166 (10)	0.39
Weight-for-age z-score, mean (SD)	-0.17 (0.98)	-0.20 (0.93)	-0.29 (1.1)	0.34
Pulmonary exacerbations in year prior to enrollment, median (range)	1 (0–7)	1 (0–6)	1 (0–14)	0.55
Pulmonary exacerbations in year prior to enrollment, mean (SD)	1.7 (1.7)	1 (1.2)	1.1 (1.5)	> 0.99
Patients with no pulmonary exacerbations, <i>n</i> (fraction affected)	79 (0.69)	63 (0.56)	7054 (0.52)	0.46
Diabetes, <i>n</i> (fraction affected)	25 (0.22)	26 (0.23)	4423 (0.31)	0.073
Pancreatic Sufficiency, <i>n</i> (fraction affected)	9 (0.079)	19 (0.17)	2239 (0.16)	0.84
CF related arthropathy, <i>n</i> (fraction affected)	8 (0.07)	10 (0.088)	697 (0.048)	0.085
5-Year Predicted Survival, median (range)	0.959 (0.094 to > 0.999)	0.973 (0.464–0.999)	0.967 (0.0949–0.999)	0.17
Home Altitude, m, mean (SD)	1305 (442)	–	–	–
Infections Present, <i>n</i> (fraction affected)				
Methicillin Sensitive <i>S aureus</i>	51 (0.45)	67 (0.59)	7937 (0.55)	0.5
Methicillin Resistant <i>S aureus</i>	21 (0.18)	33 (0.29)	5001 (0.35)	0.23
<i>P aeruginosa</i>	70 (0.61)	84 (0.74)	10,096 (0.70)	0.47
<i>B cepacia</i> complex	3 (0.026)	5 (0.044)	782 (0.054)	0.78
<i>S maltophilia</i>	7 (0.061)	26 (0.23)	2970 (0.21)	0.65
<i>Achromobacter</i> spp	5 (0.044)	11 (0.096)	1579 (0.11)	0.76
<i>Candida</i> spp	17 (0.15)	17 (0.15)	3100 (0.22)	0.11
<i>Aspergillus</i>	12 (0.11)	26 (0.23)	3091 (0.21)	0.82
<i>Mycobacterium avium-intracellulare</i> complex ^a	1 (0.0088)	8 (0.082)	730 (0.069)	0.74
<i>Mycobacterium abscessus</i> ^a	3 (0.027)	5 (0.052)	527 (0.05)	> 0.99
Treatments in use, <i>n</i> (fraction affected)				
Any Form of Inhaled Tobramycin	38 (0.33)	71 (0.62)	8069 (0.58)	0.38
Inhaled Aztreonam	40 (0.35)	47 (0.41)	4977 (0.36)	0.25
Oral Azithromycin	62 (0.54)	67 (0.59)	8595 (0.60)	0.91
Inhaled Hypertonic Saline	71 (0.62)	82 (0.72)	9611 (0.69)	0.54
Inhaled DNase	101 (0.89)	106 (0.93)	12,168 (0.87)	0.087

^aFractions reported reflect that 97 MWCFC and 10,618 CFFPR patients had acid fast bacterial cultures performed in 2014. MWCFC patients were more likely to undergo acid fast cultures than non-study patients ($p = 0.008$, χ -square test)

for HMGB-1 measurements (Additional files 3 and 4). Overnight delay in processing was associated with large increases in measured concentration for low values found after same-day sputum processing. Conversely, we found decreased values for samples processed after overnight shipping on ice that had high measurement results from same day processing. Together these effects appeared to compress the range of values, potentially reducing our ability to discriminate between patients with similar levels of inflammation (Fig. 1). These observations coupled with our intention of reproducing prior findings derived from same-day processing of sputum compelled us to require on-site processing for the current study. Shipping and

central processing of sputum might similarly affect other potential biomarker measurements, but we did not specifically address this question.

Our method of sputum sample processing repeats the method we used to discover associations between HMGB-1 and pulmonary exacerbations and survival [25]. In that effort, we modified the standard operating procedures of the Therapeutic Development Network [40] by avoiding the use of dithiothreitol (DTT) in order to optimize detection of some inflammatory biomarkers [53]. Others studies of the effects of avoiding DTT and adding protease inhibitors found that biomarker detection was improved in both cases [53–56].



Despite the care taken, our study has limitations. We originally planned to enroll up to 175 patients (Table 2, row 3) based on the analysis of numbers of patients needed to replicate our prior results with GMCSF with a high level of precision to answer our research question (Additional files 1 and 4) [25] about the relationship between GMCSF and acute FEV₁% drop with onset of a pulmonary exacerbation, a secondary goal. Unfortunately, we fell short of the planned number; however, the actual number of patients enrolled still allows us to address this question (Table 2, row 4) and well exceeded the estimated numbers of adult and pediatric patients to replicate prior results that revealed relationships between HMGB-1 and pulmonary exacerbations, our primary goal. We additionally achieved recruitment

numbers to provide sufficient power to analyze secondary outcomes including relationships between HMGB-1 and NE with pulmonary exacerbations, subsequent lung function and combined lung transplant and death outcomes (Table 2, rows 1, 2 and 5).

Because the study was non-interventional, we did not include the final question of the PHQ-9 regarding specific plans to commit suicide, because positive answers would have required intervention and a different regulatory approval pathway. Instead, we alerted clinicians to patients with high PHQ-8 scores and requested them to consider intervention as part of usual clinical care; this plan was specifically approved by all site IRB's.

We attempted to reduce bias against enrollment of patients that frequently miss clinic appointments. However,

Table 5 Collections

Collection Type	Enrollment	First Pulmonary Exacerbation Onset ^a	First Pulmonary Exacerbation Convalescence	Additional Exacerbation	End of Study Follow Up ^b
Clinical Data	114	92	36	10	72
Samples	114 ^c	52	29	8	62

^aFollow up varied and sometimes exceeded one year to the first exacerbation. Among enrolled patients, 81% had an exacerbation during the study. However, the observed percentage of patients with exacerbations within 1 year was lower, 47%, and was similar to the 44% reported in annualized 2014 CFFPR data for this cohort of patients (Table 4)

^bQueries for data from the end-of-study are ongoing at the time of submission

^c114 samples were collected, however, only 112 were sufficient in size to allow laboratory analyses

the nature of clinic non-adherence likely still reduced enrollment of these patients and may have reduced our ability to collect subsequent samples and data for our secondary study goals.

We successfully collected samples for only 52 out of 92 patients who suffered pulmonary exacerbations during study follow up. The high degree of missingness suggests that there may be hidden biases in studies of pulmonary exacerbation. Because most samples were missed due to diagnosis at non-participating care centers, distance-to-center or other reasons that impede care at an accredited CF Care Center may be previously unrecognized sources of bias in studies of pulmonary exacerbation. Further analyses of the current data and additional future studies designed to collect sputum from outlying care centers may help to understand the nature, size and effect of such bias.

Our study did not enroll patients with the same clinical status as in prior studies [5, 24, 25, 27]. Prior patients were non-randomly selected often after or during a pulmonary exacerbation. We reasoned that predictions from biomarkers measured during clinical stability would be most generally useful, thus our current study patients may be less ill as a group than previously reported patients.

Conclusions

We designed and implemented an observational study of stable patients with CF able to produce sputum to identify inflammatory biomarkers predictive of clinically important outcomes. Our design and enrollment efforts recruited a well characterized cohort from the MWCFC region similar to sputum-producing patients throughout the US that participate in the CFFPR. This cohort provides a believable and generalizable foundation in sufficient numbers for clinical interpretation of upcoming analyses of biochemical marker measurements from the carefully collected and annotated sputum samples.

Additional files

Additional file 1: Part 1 of the Video: Principles of Study Design includes overviews of Statistics and its meaning, the three main types of study design, and six principles that should guide design of studies. Total run time: 10 minutes 14 seconds. (M4V 342000 kb)

Additional file 2: Part 2 of the Video: Principles of Study Design includes a more in depth discussion of four of the six guiding principles of study design. Sir David and Ruth discuss that (1) studies should address interesting questions, (2) there should be a population of patients appropriate for the study questions, (3) measurements within the population should be done well, and (4) studies should avoid systematic error by including techniques such as balancing of patients and their characteristics and randomization with concealment to reduce observer bias in enrollment. Total run time: 8 minutes 17 seconds. (M4V 291000 kb)

Additional file 3: Part 3 of the Video: Principles of Study Design includes the two remaining points on study design. Sir David and Ruth discuss (5) that the number of individuals in a study should be sufficient

to get good answers to the questions and (6) that the data collected should be capable of analysis and interpretation. The video concludes with a few remarks about unexpected problems during study performance and advice about the important aspects of study design. Total run time: 8 minutes 6 seconds. (M4V 292000 kb)

Additional file 4: Transcript of Video: Principles of Study Design: A Conversation Between David Cox and Ruth Keogh. (PDF 124 kb)

Additional file 5: Final Study Protocol. Created and distributed 13 May 2013. (PDF 290 kb)

Additional file 6: Sputum Processing Instructions. Provides the laboratory flow sheet for sample collection and illustrated instructions. (PDF 706 kb)

Additional file 7: Video: Cystic Fibrosis Biomarker Study: Sputum Collection and Processing. The research team at the University of Utah demonstrate supplies and equipment needed to collect and process sputum specimens for the biomarker study. Total run time: 10 minutes 52 seconds. (M4V 105000 kb)

Abbreviations

CF: Cystic Fibrosis; CFF: Cystic Fibrosis Foundation; CFFPR: Cystic Fibrosis Foundation Patient Registry; ELISA: Enzyme Linked Immunosorbent Assay; FEV₁: Forced Expiratory Volume in one second; FEV₁%: Percent Predicted FEV₁; GAD: Generalized Anxiety Disorder; GMCSF: Granulocyte-Macrophage Colony Stimulating Factor; HBSS: Hanks Buffered Saline Solution; HMGB-1: High Mobility Group Box-1 protein; IL: Interleukin; IRB: Investigational Review Board; MCQ: Munich Chronotype Questionnaire; MRSA: Methicillin Resistant *Staphylococcus aureus*; MWCFC: Mountain West Cystic Fibrosis Consortium; NE: Neutrophil Elastase; PHQ: Patient Health Questionnaire; REDCap: Research Electronic Data Capture; SD: Standard Deviation

Acknowledgments

The authors would like to thank the Cystic Fibrosis Foundation for the use of CF Foundation Patient Registry data to conduct this study. Additionally, we would like to thank the patients, care providers, and clinic coordinators at CF centers throughout the United States for their contributions to the CF Foundation Patient Registry. We thank Bruce Marshall at the CF Foundation for support throughout the planning and execution of the study and for the key suggestion to recruit our External Advisory Committee.

The data and results contained in this work were presented in preliminary abstract form at the North American CF Conference in Orlando, Florida, USA, 25-28 October 2016. Preliminary, and partial study results were presented to the Pulmonary Groups at the University of Massachusetts, Worcester, MA, 25 July 2014 and the University of Virginia, Charlottesville, VA, 11 December 2014, and to the CF group at the Royal Brompton Hospital, London, UK, 20 April 2018 and at the Mountain West CF Consortium meetings 30 April 2016 and 5 May 2018. The results, interpretations and opinions expressed in this work are those of the authors and do not necessarily represent the views of the NIH, NHLBI, NCATS, NSF, the Veteran's Administration or the Margolis Foundation, the Claudia Ruth Goodrich Stevens Family, the US CF Foundation, the European CF Society or the Medical Research Council in the UK or any of the sponsors of clinical trials that provided other support for individual authors or the governments of the US or the UK.

Funding

This project was supported by the CF Foundation (CFF) (LI0U13A0, LI0U14Y4), the National Center for Advancing Translational Science at the National Institutes of Health (NCATS/NIH 8UL1TR000105 [formerly UL1RR025764]), the Ben B and Iris M Margolis Foundation of Utah and the Claudia Ruth Goodrich Stevens Endowment Fund. Neither the sponsors of the project nor any of the sources of other support for the authors had a direct role in the development and conduct of the study. None of the sponsors of clinical trials mentioned in the competing interests section participated in any way with this trial. Prior to submission, the CF Foundation reviewed the manuscript for appropriate use of the CFFPR to protect health information.

Availability of data and materials

The data sets generated and analyzed during the current study specific to the MWCFC are not publicly available due to the ongoing nature of the study and analyses. Non-patient identifying study data will be available from the corresponding author on reasonable request as additional studies based

on the biomarker data are published. Samples are not available as they are under active investigation and analysis. They will be available by reasonable request to the corresponding author as the planned additional studies are completed and published. Concerning the CFFPR, data are available upon request through the CFF Patient Registry Comparative Effectiveness Research Committee. You can contact the committee at datarequests@cff.org. Restrictions on access to data are to ensure patient privacy for all persons in the CFF Patient Registry. Old annual reports, such as cited [34], may be requested from the CFF if no longer available on the world-wide web.

Authors' contributions

TGL is the guarantor of the manuscript and takes responsibility for the integrity of the work as a whole from inception to published article. TGL drafted the initial manuscript, and all authors contributed by editing and correcting the final manuscript. TGL, FRA, RK, DRC developed the statistical plan for enrollment and the randomized selection scheme (with input from JLJ, KRP and SDS) and performed the power analyses and the statistical analysis of results. JLJ oversaw IRB submissions throughout the MWCFC with help at each site from NA, DD, JAF, BG, CMK, OM, KAP, KRP, AJR, JBV and RY. TGL, NA, PSB, BAC, CLD, DD, JAF, BG, JLJ, TH, JRH, CMK, YL, JL, NL, OM, CN, KAP, KRP, ALQ, PR, AJR, SDS, JLT-C, JBV, RY, JPC, JSE, KNO developed and performed the clinical processes for enrollment of patients and collections of samples and clinical annotations. JLJ oversaw the work leading up to the general study planning meeting with critical assistance from KAP, JAF and JBV. JLJ oversaw and coordinated the logistics of the entire study. RK and DRC prepared Additional files 1, 2 and 3, Video: Principles of Study Design, at the London School of Hygiene and Tropical Medicine 29 March 2013 for presentation at the MWCFC study meetings 11–13 April 2013. TGL prepared the transcript in Additional file 4. TGL, JAF, JLJ, YL, KAP and JBV developed and tested the sputum processing protocol and performed the shipping controls experiment. JAF, JLJ, KAP, YL and JBV provided central and on-site training of laboratory procedures and prepared Additional file 7, Video: Cystic Fibrosis Biomarker Study: Sputum Collection and Processing. TGL and JLJ performed site initiation visits. JLJ and KAP monitored data collection centrally, and JAF, KAP and JBV performed interim and end-of-study site visits to verify data integrity and compliance with regulatory requirements. KAP developed and implemented REDCap electronic data capture forms with input from TGL, JAF, JLJ and JBV. KAP oversaw REDCap data entry, query generation and resolution and initial generation of raw data reports. JBV with assistance from JAF, JLJ and YL developed and oversaw shipping protocols for all specimens by NA, DD, BG, CMK, OM, KRP, AJR and RY. JAF, JLJ, JBV, KAP and YL received, handled and successfully managed sputum specimens centrally despite shipment damage, floods, regular and emergency power outages, freezer and complete building failures. TGL obtained funding with assistance from FRA, JLJ and SDS and additional help from PSB, BAC, CLD, RK, JL, CN and DRC. JPC, JSE and KNO were the members of the Advisory Committee for the study and contributed to multiple aspects of the study. All authors had access to the complete data set from this study. All authors read and approved the final manuscript.

Authors' information

DRC is an Honorary Fellow at Nuffield College. His pioneering work in Statistics includes introduction of the current methods of logistic regression [57] and proportional hazards models [58]. He was awarded the Guy Medal in both Silver and Gold from the Royal Statistical Society, the Kettering Prize and Gold Medal for Cancer Research, the Copley Medal from the Royal Society, and he recently won the first International Prize in Statistics awarded jointly by the American Statistical Association, the International Biometric Society, the Institute of Mathematical Statistics, the International Statistical Institute, and the Royal Statistical Society. His lifelong experience improving the design of experiments [31, 32] was instrumental in the design of the current study.

RK received her DPhil under the supervision of DRC from the University of Oxford. She is now Associate Professor in the Department of Medical Statistics at the London School of Hygiene and Tropical Medicine, is funded by a Medical Research Council Methodology Fellowship, and teaches Master's level courses in medical statistics. She helped the study greatly by participating in the teaching video (Additional files 1, 2 and 3) with DRC and with our initial study design.

Our non-MWCFC collaborators brought skills and experience not found in the MWCFC. NL is Director of the Adult CF Program and Associate Professor in the Department of Medicine, Division of Pulmonary and Critical Care

Medicine at the Johns Hopkins University. He has published in the area of assessment of pain in patients with CF [45] and brings that expertise to our study. ALQ is a behavioral scientist and clinical psychologist at the Miami Children's Research Institute. She was awarded the inaugural Mattingly Leadership in Mental Health Care Award in 2016 by the US CF Foundation, and she contributed her expertise on quantitative measurements of depression, anxiety and quality of life [59] to the study.

The study benefited greatly from guidance by the external advisory committee. JPC is Professor and Research Director for the Division of Pulmonary Medicine, Department of Pediatrics, at the U of Cincinnati and has been a member or leader of multiple research organizations within the CF Foundation, the European CF Society and the National Institutes of Health in the US (NIH). KNO is Senior Clinician and Chief of the Pulmonary Branch of the National Heart, Lung and Blood Institute (NHLBI) of the NIH and leader of the Chronic Airway Infection Laboratory at the NHLBI. He has led multiple studies of bronchiectasis and infections, especially non-tuberculous *Mycobacteria*. JSE is Faculty Pro-Vice Chancellor of the School of Medicine, Dentistry and Biomedical Sciences and Faculty in the Institute for Health Sciences and the Centre for Experimental Medicine, Queen's University Belfast, UK. He was President of the European CF Society for 8 years and Trustee, Chair or Member of the Medical Advisory Committee of the CF Trust in the UK since 2002. He received knighthood in 2012 for his work in health care in Northern Ireland. He has been author and driving force for many studies that have improved knowledge and care of patients with CF. The authors from the Mountain West CF Consortium have been meeting and working collaboratively for over 20 years including a recent publication [52]. The long-standing close relationships between these individuals and teams enhanced the ability to design and perform this complex multicenter study.

Ethics approval and consent to participate

Our project was reviewed and approved by the Investigational Review Board (IRB) of the University of Utah (IRB_00011571), the St. Luke's Health System IRB (IRB Protocol No: 13-0547), the IRB of Billings (Study 14.07), the Western IRB (for the Las Vegas CF Center) (STUDY NUM 1146159, WIRB PRO NUM: 20140721, INVEST NUM: 112264, WO NUM: 1-837046-1, Protocol Number 20130530), the National Jewish Health IRB (Study Number: HS-2797), Colorado Multiple IRB (for Children's Hospital Colorado) (COMIRB Protocol 13-3172), the Phoenix Children's Hospital IRB (PCH IRB #13-087), the University of Arizona IRB (Protocol Number: 1407394040), and the Human Research Review Committee in the Human Research Protections Office at the University of New Mexico (Study ID 14-085). No patients were enrolled prior to IRB approval. All patients younger than 18 provided assent with informed written consent provided by parents or other legal guardians. Each adult patient gave informed written consent on his or her own behalf.

Consent for publication

Not applicable.

Competing interests

TGL, JAF, JLJ, YL, KAP and JBV received other support from the CFF (CC132-16 AD, LIOU14Y0, LIOU14P0) and the National Heart Lung and Blood Institute (NHLBI) of the National Institutes of Health (NIH) (R01 HL125520) and received support during the current study for performing clinical trials from Abbvie, Gilead, Nivalis, Novartis, Proteostasis, Inc., Savara and Vertex. FRA received additional other support from the NHLBI/NIH (R01 HL125520), the National Science Foundation (EMSW21-RTG) and the Margolis Foundation of Utah. PSB received other support from the CFF (Center and TDC grants) and the NHLBI/NIH (U01 HL114623) and received support for a clinical trial from Alcresta. BAC received other support from the CFF (C112-12, C112-TDC09Y, 10063SUB, 41,339,154.s132P010379SUB) and received support for clinical trials from Genentech, Novartis and Vertex. CLD received other support from the CFF (C004-11, C004-TDC09Y, DAINES11Y3) and from the Health Resources and Services Administration (T72MC00012). TH received other support from the CFF (PACE, Center Grant) and received support for clinical trials from Celtaxsys and Vertex. JRH received other support from the NHLBI/NIH (HHSN268200900018C) and the Veterans Administration Healthcare System (I01 BX001533). JL received other support from the CFF (C017-11AF). CN received other support from the CFF (C138-12). PR received other support from the CFF (C003-12, C003-TDC09Y). SDS received other support from the CFF (AQUADEK12K1, SAGEL11CS0, GOAL13K2, NICK13A0, SAGEL14K1, NICK15R0) and the NHLBI/NIH (U54 HL096458) and the NCATS/NIH (Colorado CTS Grant Number UL1

TR002535). JLT-C received other support from the CFF (TDC) and the NHLBI/NIH (HL103801) and received support for clinical trials from Vertex. KNO is funded by the intramural research program of the NHLBI, NIH.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Adult Cystic Fibrosis Center, Division of Respiratory, Critical Care and Occupational Pulmonary Medicine, Department of Internal Medicine, University of Utah, 26 North Mario Capecchi Drive, Salt Lake City, UT 84132, USA. ²Intermountain Pediatric Cystic Fibrosis Center, Division of Pediatric Pulmonology, Department of Pediatrics, University of Utah, 81 North Mario Capecchi Drive, Salt Lake City, UT 84113, USA. ³Departments of Mathematics, University of Utah, 155 South 1400 east, JWB 233, Salt Lake City, UT 84112, USA. ⁴School of Biological Sciences, University of Utah, 257 South 1400 East, Salt Lake City, UT 84112, USA. ⁵Cystic Fibrosis Center, Phoenix Children's Hospital, 1919 East Thomas Road, Phoenix, AZ 85016, USA. ⁶St. Luke's Cystic Fibrosis Center of Idaho, 610 W. Hays Street, Boise, ID 83702, USA. ⁷Division of Pediatric Pulmonary and Sleep Medicine, Department of Pediatrics, University of Arizona Health Sciences, 1501 N. Campbell Avenue, Room 3301, PO Box 245073, Tucson, AZ 85724, USA. ⁸Cystic Fibrosis Center, 3006 S. Maryland Pkwy, Suite #315, Las Vegas, NV 89109, USA. ⁹Adult Cystic Fibrosis Program, Division of Pulmonary, Critical Care and Sleep Medicine, DoIM MSC10-5550, 1 University of New Mexico, Albuquerque, NM 87131, USA. ¹⁰Department of Medical Statistics, London School of Hygiene and Tropical Medicine, Room G36, Keppel Street, London WC1E 7HT, UK. ¹¹Department of Pediatrics, Children's Hospital Colorado, University of Colorado School of Medicine, 13123 East 16th Avenue, Aurora, CO 80045, USA. ¹²Division of Pulmonary and Critical Care and Sleep Medicine, Department of Medicine, Johns Hopkins University School of Medicine, 601 N. Caroline St, Baltimore, MD 21287, USA. ¹³Montana Cystic Fibrosis Center, Billings Clinic, 2800 10th Avenue N, Billings, MT 59101, USA. ¹⁴Division of Pulmonary and Critical Care and Sleep Medicine, Department of Medicine, National Jewish Health, 1400 Jackson Street, Denver, CO 80206, USA. ¹⁵Former: Department of Psychology, University of Miami, Miami, FL, USA. ¹⁶Present Address: Miami Children's Research Institute, Nicklaus Children's Hospital, 3100 SW 62nd Ave, Miami, FL 33155, USA. ¹⁷Division of Pulmonology, Department of Pediatrics, National Jewish Health, 1400 Jackson St, Denver, CO 80206, USA. ¹⁸Division of Pulmonary Medicine, Department of Pediatrics, University of Cincinnati, 3333 Burnet Avenue, Cincinnati, OH 45229-3026, USA. ¹⁹Faculty of Medicine, Health and Life Sciences, Queen's University Belfast, 90 Lisburn Road, Belfast BT9 6AG, UK. ²⁰Laboratory of Chronic Airway Infection, Pulmonary Branch, National Heart Lung and Blood Institute, National Institutes of Health, 10 Center Drive MSC1454, Building 10-CRC, Room 1408A, Bethesda, MD 20892, USA. ²¹Nuffield College, 1 New Rd, Oxford OX1 1NF, UK.

Received: 18 September 2018 Accepted: 11 March 2019

Published online: 26 April 2019

References

- Gibson RL, Burns JL, Ramsey BW. Pathophysiology and management of pulmonary infections in cystic fibrosis. *Am J Respir Crit Care Med.* 2003;168(8):918–51.
- Sagel SD, Chmiel JF, Konstan MW. Sputum biomarkers of inflammation in cystic fibrosis lung disease. *Proc Am Thorac Soc.* 2007;4(4):406–17.
- Mayer-Hamblett N, Aitken ML, Accurso FJ, et al. Association between pulmonary function and sputum biomarkers in cystic fibrosis. *Am J Respir Crit Care Med.* 2007;175(8):822–8.
- Bonfield TL, Panuska JR, Konstan MW, et al. Inflammatory cytokines in cystic fibrosis lungs. *Am J Respir Crit Care Med.* 1995;152(6 Pt 1):2111–8.
- Sagel SD, Wagner BD, Anthony MM, Emmett P, Zemanick ET. Sputum biomarkers of inflammation and lung function decline in children with cystic fibrosis. *Am J Respir Crit Care Med.* 2012;186:857–65.
- Ramsey BW, Banks-Schlegel S, Accurso FJ, et al. Future directions in early cystic fibrosis lung disease research: an NHLBI workshop report. *Am J Respir Crit Care Med.* 2012;185(8):887–92.
- Levy H, Kalish LA, Huntington I, et al. Inflammatory markers of lung disease in adult patients with cystic fibrosis. *Pediatr Pulmonol.* 2007;42(3):256–62.
- Saavedra M, Hughes G, Sanders L, et al. Circulating RNA transcripts identify therapeutic response in cystic fibrosis lung disease. *Am J Respir Crit Care Med.* 2008;178(9):929–38.
- Liou TG, Adler FR, Fitzsimmons SC, Cahill BC, Hibbs JR, Marshall BC. Predictive 5-year survivorship model of cystic fibrosis. *Am J Epidemiol.* 2001;153(4):345–52.
- Buzzetti R, Alicandro G, Minicucci L, et al. Validation of a predictive survival model in Italian patients with cystic fibrosis. *J Cyst Fibros Off J Eur Cyst Fibros Soc.* 2012;11(1):24–9.
- Fuchs HJ, Borowitz DS, Christiansen DH, et al. Effect of aerosolized recombinant human DNase on exacerbations of respiratory symptoms and on pulmonary function in patients with cystic fibrosis. *N Engl J Med.* 1994;331(10):637–42.
- Ramsey BW, Pepe MS, Quan JM, et al. Intermittent Administration of Inhaled Tobramycin in patients with cystic fibrosis. *N Engl J Med.* 1999;340(1):23–30.
- Saiman L, Marshall BC, Mayer-Hamblett N, et al. Azithromycin in patients with cystic fibrosis chronically infected with *Pseudomonas aeruginosa*: a randomized controlled trial. *JAMA J Am Med Assoc.* 2003;290(13):1749–56.
- Elkins MR, Robinson M, Rose BR, et al. A controlled trial of long-term inhaled hypertonic saline in patients with cystic fibrosis. *N Engl J Med.* 2006;354(3):229–40.
- Retsch-Bogart GZ, Quittner AL, Gibson RL, et al. Efficacy and safety of inhaled aztreonam lysine for airway *Pseudomonas* in cystic fibrosis. *Chest.* 2009;135(5):1223–32.
- Ramsey BW, Davies J, McElvaney NG, et al. A CFTR potentiator in patients with cystic fibrosis and the G551D mutation. *N Engl J Med.* 2011;365(18):1663–72.
- Boyle MP, Bell SC, Konstan MW, et al. A CFTR corrector (lumacaftor) and a CFTR potentiator (ivacaftor) for treatment of patients with cystic fibrosis who have a phe508del CFTR mutation: a phase 2 randomised controlled trial. *Lancet Respir Med.* 2014;2(7):527–38.
- Saiman L, Anstead M, Mayer-Hamblett N, et al. Effect of azithromycin on pulmonary function in patients with cystic fibrosis uninfected with *Pseudomonas aeruginosa*: a randomized controlled trial. *JAMA.* 2010;303(17):1707–15.
- Fidler KJ, Hilliard TN, Bush A, et al. Mannose-binding lectin is present in the infected airway: a possible pulmonary defence mechanism. *Thorax.* 2009;64(2):150–5.
- Hubeau C, Le Naour R, Abély M, et al. Dysregulation of IL-2 and IL-8 production in circulating T lymphocytes from young cystic fibrosis patients. *Clin Exp Immunol.* 2004;135(3):528–34.
- Koller DY, Nething J, Otto J, Urbanek R, Eichler I. Cytokine concentrations in sputum from patients with cystic fibrosis and their relation to eosinophil activity. *Am J Respir Crit Care Med.* 1997;155(3):1050–4.
- Merenmies J, Pihlaskari R, Laitinen J, Wartiovaara J, Rauvala H. 30-kDa heparin-binding protein of brain (amphoterin) involved in neurite outgrowth. Amino acid sequence and localization in the filopodia of the advancing plasma membrane. *J Biol Chem.* 1991;266(25):16722–9.
- Wang H, Bloom O, Zhang M, et al. HMG-1 as a late mediator of endotoxin lethality in mice. *Science.* 1999;285(5425):248–51.
- Rowe SM, Jackson PL, Liu G, et al. Potential role of high-mobility group box 1 in cystic fibrosis airway disease. *Am J Respir Crit Care Med.* 2008;178(8):822–31.
- Liou TG, Adler FR, Keogh RH, et al. Sputum biomarkers and the prediction of clinical outcomes in patients with cystic fibrosis. *PLoS One.* 2012;7(8):e42748.
- Gray RD, MacGregor G, Noble D, et al. Sputum proteomics in inflammatory and suppurative respiratory diseases. *Am J Respir Crit Care Med.* 2008;178(5):444–52.
- Gray RD, Imrie M, Boyd AC, Porteous D, Innes JA, Greening AP. Sputum and serum calprotectin are useful biomarkers during CF exacerbation. *J Cyst Fibros Off J Eur Cyst Fibros Soc.* 2010;9(3):193–8.
- Goldstein W, Döring G. Lysosomal enzymes from polymorphonuclear leukocytes and proteinase inhibitors in patients with cystic fibrosis. *Am Rev Respir Dis.* 1986;134(1):49–56.
- Nakamura H, Yoshimura K, McElvaney NG, Crystal RG. Neutrophil elastase in respiratory epithelial lining fluid of individuals with cystic fibrosis induces interleukin-8 gene expression in a human bronchial epithelial cell line. *J Clin Invest.* 1992;89(5):1478–84.
- Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med.* 1999;159(1):179–87.

31. Cox DR. Planning of experiments. New York: Wiley; 1958.
32. Cox DR, Donnelly CA. Principles of Applied Statistics. 1st ed. Cambridge: Cambridge University Press; 2011.
33. 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):1173–9.
34. Cystic Fibrosis Foundation Patient Registry 2016 Annual Data Report [Internet]. Bethesda, Maryland: Cystic Fibrosis Foundation; 2017. Available from: <https://www.cfsmartreports.com/pages/prelogin/about>. Accessed 27 Mar 2019.
35. Geography UCB. Geography Reference [Internet]. Available from: <https://www.census.gov/geo/reference/state-area.html>. Accessed 27 Mar 2019.
36. Rosenfeld M, Emerson J, Williams-Warren J, et al. Defining a pulmonary exacerbation in cystic fibrosis. *J Pediatr*. 2001;139(3):359–65.
37. Davison AC. Bootstrap Methods And Their Application. 1 edition. Cambridge; New York, NY, USA: Cambridge University Press; 1997.
38. Davé SH, Tilstra JS, Matsuoka K, et al. Ethyl pyruvate decreases HMGB1 release and ameliorates murine colitis. *J Leukoc Biol*. 2009;86(3):633–43.
39. Mogayzel PJ, Naureckas ET, Robinson KA, et al. Cystic fibrosis pulmonary guidelines. Chronic medications for maintenance of lung health. *Am J Respir Crit Care Med*. 2013;187(7):680–9.
40. Hilliard JB, Konstan MW, Davis PB. Inflammatory mediators in CF patients. *Methods Mol Med*. 2002;70:409–31.
41. Borg GA. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc*. 1982;14(5):377–81.
42. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. 2006;166(10):1092–7.
43. Quittner AL, Goldbeck L, Abbott J, et al. Prevalence of depression and anxiety in patients with cystic fibrosis and parent caregivers: results of the international depression epidemiological study across nine countries. *Thorax*. 2014;69(12):1090–7.
44. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. *J Gen Intern Med*. 2001;16(9):606–13.
45. Hayes M, Yaster M, Haythornthwaite JA, et al. Pain is a common problem affecting clinical outcomes in adults with cystic fibrosis. *Chest*. 2011;140(6):1598–603.
46. Roenneberg T, Wirz-Justice A, Meroo M. Life between clocks: daily temporal patterns of human Chronotypes. *J Biol Rhythm*. 2003;18(1):80–90.
47. Nord M. Assessing Potential Technical Enhancements to the U.S. Household Food Security Measures [Internet]. US Department of Agriculture, Economic Research Service; 2012 [cited 2016 Jul 19]. Available from: <https://www.ers.usda.gov/publications/pub-details/?pubid=47606>. Accessed 27 Mar 2019.
48. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377–81.
49. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2015. Available from: <https://www.r-project.org/>. Accessed 27 Mar 2019.
50. TNM Download [Internet]. [cited 2016 Jul 19]; Available from: <https://viewer.nationalmap.gov/basic/?basemap=b1&category=ned,nedsrc&title=3DEP%20View>. Accessed 27 Mar 2019.
51. Quanjer PH, Stanojevic S, Cole TJ, 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):1324–43.
52. Liou TG, Jensen JL, Allen SE, et al. Improving performance in the detection and management of cystic fibrosis-related diabetes in the mountain west cystic fibrosis consortium. *BMJ Open Diabetes Res Care*. 2016;4(1):e000183.
53. Kelly MM, Keatings V, Leigh R, et al. Analysis of fluid-phase mediators. *Eur Respir J Suppl*. 2002;37:24s–39s.
54. Smountas AA, Lands LC, Mohammed SR, Grey V. Induced sputum in cystic fibrosis: within-week reproducibility of inflammatory markers. *Clin Biochem*. 2004;37(11):1031–6.
55. Hector A, Jonas F, Kappler M, Feilcke M, Hartl D, Griese M. Novel method to process cystic fibrosis sputum for determination of oxidative state. *Respir Int Rev Thorac Dis*. 2010;80(5):393–400.
56. Kelly MM, Leigh R, Carruthers S, et al. Increased detection of interleukin-5 in sputum by addition of protease inhibitors. *Eur Respir J Off J Eur Soc Clin Respir Physiol*. 2001;18(4):685–91.
57. Cox DR. The regression analysis of binary sequences. *J R Stat Soc Ser B Methodol*. 1958;20(2):215–42.
58. Cox DR. Regression models and life-tables. *J R Stat Soc Ser B Methodol*. 1972;34(2):187–220.
59. Quittner AL, Buu A, Messer MA, Modi AC, Watrous M. Development and validation of the cystic fibrosis questionnaire in the United States: a health-related quality-of-life measure for cystic fibrosis. *Chest*. 2005;128(4):2347–54.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

