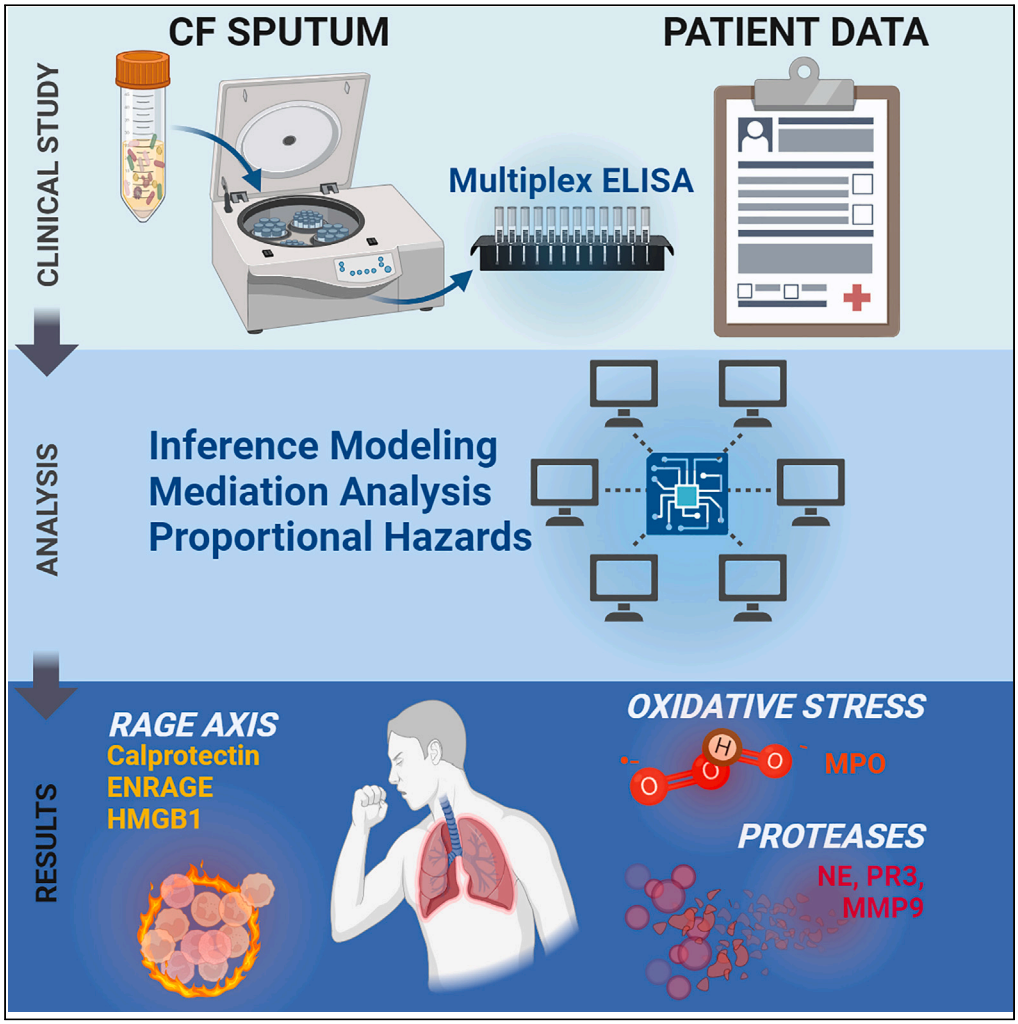
iScience

CellPress
OPEN ACCESS

Article

Airway inflammation accelerates pulmonary exacerbations in cystic fibrosis



Theodore G. Liou, Natalia Argel, Fadi Asfour, ..., J. Stuart Elborn, Kenneth N. Olivier, Frederick R. Adler

ted.liou@utah.edu

Highlights

Up to 10 biomarkers are associated with subsequent CF pulmonary exacerbations

FEV_1% mediates and obscures biomarker effects

RAGE, protease, and oxidative stress inflammation underlie CF pulmonary exacerbations

Results are generalizable to people with CF in the US in the early modulator era

Liou et al., iScience 27, 108835
March 15, 2024 © 2024 The Author(s).
https://doi.org/10.1016/j.isci.2024.108835

iScience

CellPress
OPEN ACCESS

Article

Airway inflammation accelerates pulmonary exacerbations in cystic fibrosis

Theodore G. Liou,[1,2,27,*] Natalia Argel,[3] Fadi Asfour,[2] Perry S. Brown,[4] Barbara A. Chatfield,[2] David R. Cox,[5] Cori L. Daines,[6] Dixie Durham,[5] Jessica A. Francis,[1,23] Barbara Glover,[7] My Helms,[1] Theresa Heynekamp,[8] John R. Hoidal,[1] Judy L. Jensen,[1] Christiana Kartsonaki,[9] Ruth Keogh,[10] Carol M. Kopecky,[11] Noah Lechtzin,[12] Yanping Li,[1] Jerimiah Lysinger,[13] Osmara Molina,[6] Craig Nakamura,[7] Kristyn A. Packer,[1] Robert Paine III,[1] Katie R. Poch,[14] Alexandra L. Quittner,[15] Peggy Radford,[3] Abby J. Redway,[8,24] Scott D. Sagel,[11] Rhonda D. Szczesniak,[16] Shawna Sprandel,[13] Jennifer L. Taylor-Cousar,[14,18] Jane B. Vroom,[1,2] Ryan Yoshikawa,[7] John P. Clancy,[17,25] J. Stuart Elborn,[19] Kenneth N. Olivier,[20,26] and Frederick R. Adler[21,22]

SUMMARY

Airway inflammation underlies cystic fibrosis (CF) pulmonary exacerbations. In a prospective multicenter study of randomly selected, clinically stable adolescents and adults, we assessed relationships between 24 inflammation-associated molecules and the future occurrence of CF pulmonary exacerbation using proportional hazards models. We explored relationships for potential confounding or mediation by clinical factors and assessed sensitivities to treatments including CF transmembrane regulator (CFTR) protein synthesis modulators. Results from 114 participants, including seven on ivacaftor or lumacaftor-ivacaftor, representative of the US CF population during the study period, identified 10 biomarkers associated with future exacerbations mediated by percent predicted forced expiratory volume in 1 s. The findings were not sensitive to anti-inflammatory, antibiotic, and CFTR modulator treatments. The analyses suggest that combination treatments addressing RAGE-axis inflammation, protease-mediated injury, and oxidative stress might prevent pulmonary exacerbations. Our work may apply to other airway inflammatory diseases such as bronchiectasis and the acute respiratory distress syndrome.

[1]Adult 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
[2]Primary Children's Cystic Fibrosis Center, Division of Pediatric Pulmonology, Department of Pediatrics, University of Utah, 81 North Mario Capecchi Drive, Salt Lake City, UT 84113, USA
[3]Cystic Fibrosis Center, Phoenix Children's Hospital, 1919 East Thomas Road, Phoenix, AZ 85016, USA
[4]St. Luke's Cystic Fibrosis Center of Idaho, 610 W. Hays Street, Boise, ID 83702, USA
[5]Nuffield College, 1 New Rd, Oxford OX1 1NF, UK
[6]Division of Pediatric Pulmonary and Sleep Medicine, Department of Pediatrics, University of Arizona Health Sciences, University of Arizona, 1501 N. Campbell Avenue, Room 3301, PO Box 245073, Tucson, AZ 85724, USA
[7]Cystic Fibrosis Center, 3006 S. Maryland Pkwy, Suite #315, Las Vegas, NV 89109, USA
[8]Adult Cystic Fibrosis Program, Division of Pulmonary, Critical Care and Sleep Medicine, DoIM MSC10-5550, 1 University of New Mexico, Albuquerque, NM 87131, USA
[9]Clinical Trial Service Unit & Epidemiological Studies Unit and Medical Research Council Population Health Research Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK
[10]Department of Medical Statistics, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK
[11]Department of Pediatrics, Children's Hospital Colorado and University of Colorado Anschutz Medical Campus, 13123 East 16th Avenue, Aurora, CO 80045, USA
[12]Division of Pulmonary and Critical Care and Sleep Medicine, Department of Medicine, Johns Hopkins University School of Medicine, 1830 E. Monument Street, Baltimore, MD 21205, USA
[13]Montana Cystic Fibrosis Center, Billings Clinic, 2800 10th Avenue N, Billings, MT 59101, USA
[14]Division of Pulmonary and Critical Care and Sleep Medicine, Department of Medicine, National Jewish Health, 1400 Jackson Street, Denver, CO 80206, USA
[15]Behavioral Health Systems Research, Miami, FL 33133, USA
[16]Division of Biostatistics & Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA
[17]Former: Division of Pulmonary Medicine, Department of Pediatrics, University of Cincinnati, Cincinnati, OH, USA
[18]Division of Pulmonology, Department of Pediatrics, National Jewish Health, 1400 Jackson St, Denver, CO 80206, USA
[19]School of Medicine, Dentistry and Biomedical Sciences, Queen's University, Health Sciences Building, Lisburn Rd, Belfast BT9 7AE, UK
[20]Laboratory 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
[21]Department of Mathematics, 155 South 1400 East, University of Utah, Salt Lake City, UT 84112, USA
[22]School of Biological Sciences, 257 South 1400 East, University of Utah, Salt Lake City, UT 84112, USA
[23]Present address: with Vertex Pharmaceuticals
[24]Present address: with Los Alamos National Laboratory
[25]Present address: Cystic Fibrosis Foundation, 4550 Montgomery Avenue, Bethesda, MD 20814, USA
[26]Present address: Division of Pulmonary Diseases and Critical Care Medicine, 130 Mason Farm Road, CB#7020, 4th Floor Bioinformatics Building, Chapel Hill, NC 27599
[27]Lead contact
*Correspondence: ted.liou@utah.edu
https://doi.org/10.1016/j.isci.2024.108835

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

INTRODUCTION

Airway inflammation erodes lung health in cystic fibrosis (CF). Infecting organisms initiate inflammation in infants,[1] and the resulting chronic inflammation and microbial interactions[2] lead to airway obstruction, mucus impaction, microbial biofilms, and bronchiectasis. Acute events such as viral infections potentially trigger exacerbations that often lead to stepwise reductions in lung function[3] even in patients treated with elexacaftor-tezacaftor-ivacaftor combination therapy. This latest triple combination CF transmembrane regulator (CFTR) protein modulator modifies sputum characteristics and both inflammation and microbiology, but the degrees of improvements remain incompletely characterized.[4–6] There is growing agreement that better understandings of inflammation and new treatments are needed[7–10] beyond the modulator therapies in use, particularly because substantial numbers of patients cannot be treated with the newest agents due to genetic ineligibility, serious adverse events,[11] or high healthcare access barriers at individual or national levels.[12] Many patients continue to experience difficult treatment burdens and reduced quality of life, due in part to inflammation-driven exacerbations and associated hospitalizations, and ultimately suffer early mortality[13] primarily due to respiratory failure.[14–16]

We previously found that high-mobility group box 1 protein (**HMGB1**, see also Figure 1 Legend for all biomarker abbreviations) in expectorated sputum was associated with time to next exacerbation in proportional hazards modeling when adjusted by number of prior year pulmonary exacerbations.[17] Sputum calprotectin and neutrophil elastase (**NE**) were similarly associated with subsequent pulmonary exacerbations.[18,19] Other inflammatory molecules are associated in multiple settings with concurrent or past CF pulmonary exacerbations but were not tested for associations with subsequent events.[20–28]

Seeking inroads toward an integrated understanding of complex airway processes and clinical outcomes, we studied clinically stable participants with CF, sputum-derived inflammation, and subsequent pulmonary exacerbations. Nine clinical variables (age, sex, percent predicted forced expiratory volume in 1 s [**FEV_1%**], number of prior year pulmonary exacerbations, rates of pancreatic sufficiency, diabetes, and methicillin-sensitive *Staphylococcus aureus* [**MSSA**] and *Burkholderia cepacia* complex [**BCC**] infections) successfully predict survival.[14,29] These variables may be associated with underlying inflammation. Clinical abnormalities such as tachypnea or acutely decreased lung function help define pulmonary exacerbations,[13] but they may also be associated with increasing inflammation. Variables that describe clinical states that may be associated with inflammation may also be able to confound and mediate[30–32] associations between inflammatory biomarkers and future exacerbations but have received limited attention in published studies.

In this study, we sought to validate relationships between HMGB1, calprotectin, NE, and other biomarkers with pulmonary exacerbations through prospective observation. Because our observations begin during chronically ongoing infection and inflammation and developing structural lung disease, albeit during clinical stability, we included concurrent disease characteristics to help evaluate potential confounding and mediation that increase the difficulty of analyses of inflammation. The study incorporates randomized patient selection to minimize observer bias and maximize generalizability.[33] We collected spontaneously expectorated sputum to measure airway inflammation, and we used a directed acyclic graph to guide the strategy for analyses (Figure 1). We synthesized findings to better integrate our understanding of the inflammatory milieu underlying pulmonary exacerbations and widen the identification of treatment targets for further investigation in CF.

RESULTS

Participant characteristics

After Mountain West Cystic Fibrosis Consortium (MWCFC) IRB approvals (Table S1) and written informed consent, we enrolled 114 participants who were clinically stable and able to expectorate sputum. The participants enrolled exceeded pre-study estimates of numbers needed to provide sufficient power for evaluating the relationships of HMGB1 with future pulmonary exacerbations and NE with FEV_1%, and they approached the number needed, 125, for evaluation of the relationship between GMCSF and FEV_1% changes with exacerbations.[17] We previously found that the participants were broadly representative of US adolescents and adults with CF who produce sputum.[33] All participants completed follow-up to the next pulmonary exacerbation or study end, up to 2.65 years after enrollment.

Compared with participants with none, those with any pulmonary exacerbations within the year post-enrollment tended toward higher respiratory and heart rates, lower FEV_1%, more numerous pre-enrollment exacerbations, higher CF-related diabetes (**CFRD**) prevalence, and poorer five-year prognostic risk scores (Table 1). Kaplan-Meier analyses[34] showed that five-year predicted survival risk scores, prior year exacerbations, diabetes status, and respiratory rate (but not lower FEV_1%) were associated with time to next exacerbation (Figure 2). Univariable logistic regressions[35] suggested that lower FEV_1%, more frequent prior pulmonary exacerbations, CF-related diabetes, and higher heart and respiratory rates were associated with higher odds of experiencing a pulmonary exacerbation during follow-up (Table S2). Proportional hazards modeling[36] showed that younger age and lower weight-for-age Z score also increased the hazard of pulmonary exacerbation (Table S2). Altogether, these results suggested these clinical characteristics as potential confounders or mediators of biomarker effects for time to next pulmonary exacerbation.[30,31]

Biomarkers and confounding and mediation analyses

Multiple correlations between biomarkers remained significant after family-wise error rate correction for $\alpha \leq 1 \times 10^{-5}$ (Table S3). Strong correlations (magnitude ≥ 0.80) were all positive and observed between MPO and ENRAGE; interleukin-1β (IL-1β) and MMP9; sRAGE and NE; and among TARC, IL-5, and interferon gamma (IFNγ). Strong statistical correlations may indicate biologically relevant relationships.

We assessed confounding and mediation by clinical variables (Figure S1).[30,31] Intrinsic variables, age, sex, and prior pulmonary exacerbations, cannot mediate inflammation but can confound analyses.[37] We tested each one as the dependent variable for associations with each

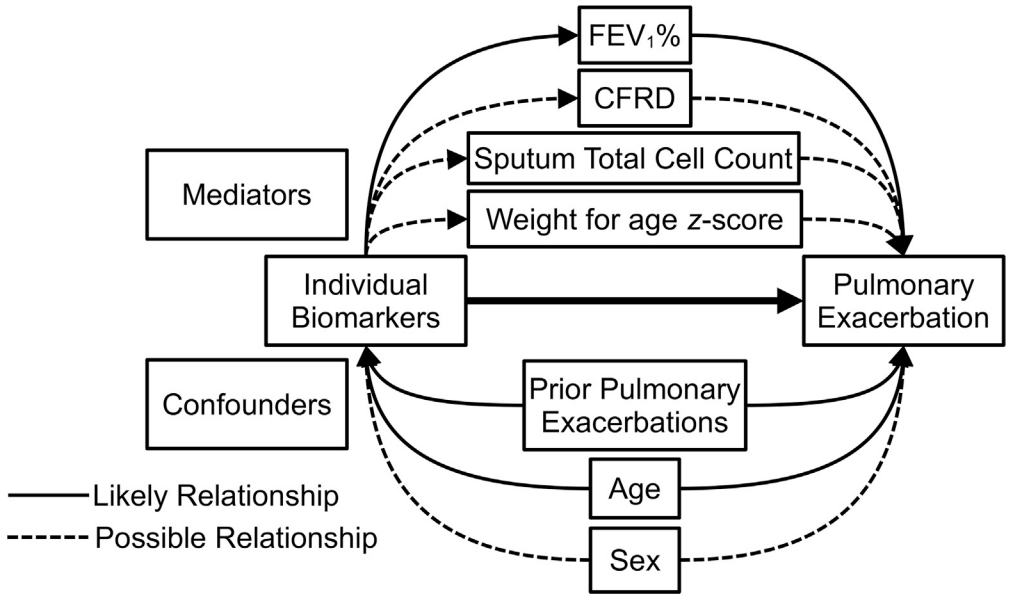


Figure 1. Directed acyclic graph of biomarkers and pulmonary exacerbation in CF

The **heavy arrow** shows the key relationships targeted by study. **Thin unidirectional arrows** indicate either **Mediator** relationships (**top**) involving variables that are intermediate outcomes in causal pathways between biomarkers and exacerbations or **Confounder** relationships (**bottom**) involving variables associated with both biomarkers and exacerbations but not as intermediates. Mediators obscure biomarker-exacerbation relationships unless excluded from explanatory models. In contrast, confounders may introduce spurious associations between biomarkers and exacerbations unless some are included as model adjustments. Not all potential confounders and mediators are shown. **Individual biomarkers** included C reactive protein (**CRP**); **calprotectin**; extracellular newly identified receptor for advanced glycation end products binding protein (**ENRAGE**); granulocyte macrophage colony stimulating factor (**GMCSF**); high-mobility group box 1 protein (**HMGB1**); intracellular adhesion molecule 1 (**ICAM1**); interferon γ (**IFNγ**); interleukin (**IL**)-**1β, -5, -6, -8, -10, -17A**; matrix metallopeptidase 9 (**MMP9**); myeloperoxidase (**MPO**); neutrophil elastase (**NE**); proteinase 3 (**PR3**); **s100A8** and **s100A9**; soluble receptor for advanced glycation end products (**sRAGE**); secretory leukoprotease inhibitor (**SLPI**); thymus and activation regulation chemokine (**TARC**); tumor necrosis factor α (**TNFα**); and chitinase 3-like 1 protein (**YKL40**).

biomarker as the independent variable using methods appropriate for each dependent variable (Table S4). Age and number of prior year pulmonary exacerbations were each associated with several biomarkers suggesting confounder roles, which we tested further by inclusion of these two variables as adjustments in proportional hazards models of time to next exacerbation (see below). Sex had no associations with any biomarker thus was not included in further models.

Univariable proportional hazards modeling showed that ENRAGE, MPO, NE, sRAGE, IL-1β, and S100A9 increased the hazard of pulmonary exacerbation. After adjusting for confounding by age and number of prior year pulmonary exacerbations, hazard ratios for ENRAGE, MPO, NE, and sRAGE increased in magnitude and significance. No models failed tests of the assumption of proportionality.[38] The large impact on the hazard ratios for these biomarkers implies that age and prior exacerbation counts act as confounders and are appropriate model adjustments. S100A9 was no longer associated with time to next pulmonary exacerbation after adjustments. Among remaining models, individual hazard ratios for YKL40, ICAM1, TARC, and MMP9 increased in magnitude and reached statistical significance at $p < 0.05$. However, because there were 24 biomarkers undergoing statistical evaluation, we performed a graphical false discovery rate analysis (FDR).[39,40] By choosing an FDR threshold < 0.2, we demonstrate that eight of the ten biomarkers with p values < 0.05 are likely to represent true findings, although the analysis cannot identify which eight (Figure 3; Table S5).

In our evaluation of mediation by different clinical factors, we found that multiple inflammatory biomarkers were strongly associated with FEV_1% (Figure 4A). Higher levels of ENRAGE, MPO, NE, sRAGE, MMP9, IL-1β, PR3, and calprotectin were associated by linear regression with lower FEV_1%, whereas SLPI and S100A8 levels were associated with higher FEV_1%. Nine associations retained significance with FDR estimates starting at < 0.001 and continuing to < 0.05 (Figures 4A and 4B). In bivariable proportional hazards models with time to next exacerbation as the outcome, none of the biomarkers retained significance ($p < 0.05$) and none had an FDR < 0.2 when modeled with FEV_1% as the second independent variable (Figure S2). After adjustment with age and number of prior year pulmonary exacerbations, p values remained greater than 0.05 (Figure 4C, red) with two exceptions for ENRAGE ($p = 0.02$) and TARC ($p = 0.04$). However, all FDRs remained above 0.2 (Figure 4D). These findings strongly indicate that FEV_1% acts as a mediator in statistical models of inflammation and should be excluded from models seeking understanding of relationships between inflammatory biomarkers and time to next exacerbation.[31]

MMP9, IL-1β, ENRAGE, IL-8, calprotectin, MPO, and PR3 were associated with log transformed sputum total cell counts (Figure S3A) and retained significance at the FDR < 0.05 level after analysis (Figure S3B). Bivariable models for time to next exacerbation using each biomarker and cell counts and adjusted by age and prior pulmonary exacerbations showed that some biomarkers retained statistical significance

Table 1. Participant characteristics and treatments for patients by exacerbations observed within one year

Characteristic or chronic treatment received	All patients n = 114	No exacerbations in first Year, n = 42	Any exacerbations in first year, n = 72	p value[a]
Years of Age[b]	26.3 (12.8–68.2)	27.8 (12.8–68.2)	24.5 (13–67.1)	0.084[d]
Respiratory Rate[b]	16 (12–30)	16 (12–26)	18 (12–30)	0.009[d]
Heart Rate[b]	82 (60–122)	80.5 (62–106)	89.5 (60–122)	0.002[d]
FEV_1%[b]	72.1 (19.6–119)	80.5 (30.2–119)	66.5 (19.6–113)	0.018[d]
Total Cell Count/mm^3, median (range)	15 (12.1–17.8)	15.2 (12.1–17.8)	14.9 (12.1–17.8)	0.20[d]
Exacerbations in Year Prior to Enrollment[b]	1 (0–7)	0 (0–5)	2 (0–7)	< 0.001[e]
Days from Last Exacerbation Prior to Enrollment[b]	183 (14–416)	216 (16–357)	182 (14–416)	0.46[d]
Weight-for-Age Z score[b]	−0.118 (−3.29 to 2.72)	−0.026 (−1.41 to 2.12)	−0.147 (−3.29 to 2.72)	0.23[d]
Female[c]	0.46	0.43	0.49	0.69[f]
Diabetes[c]	0.22	0.095	0.29	0.027[f]
Pancreatic Sufficiency[c]	0.079	0.12	0.056	0.29[f]
MSSA Infection[c]	0.45	0.48	0.43	0.78[f]
Burkholderia cepacia Complex Infection[c]	0.026	0	0.042	0.30[g]
5-Year Prognostic Risk Score[b]	3.16 (−1.8 to 6.08)	3.73 (0.676–6.08)	2.76 (−1.8 to 5.59)	< 0.001[d]
Inhaled Steroids[c]	0.59	0.5	0.64	0.21[f]
Oral Steroids[c]	0.07	0.12	0.042	0.14[f]
Any Steroids[c]	0.61	0.55	0.64	0.45[f]
Azithromycin[c]	0.54	0.48	0.58	0.36[f]
Other Oral Antibiotics[c]	0.11	0.095	0.12	0.77[f]
Any Oral Antibiotics[c]	0.61	0.5	0.67	0.12[f]
Inhaled Tobramycin[c]	0.33	0.31	0.35	0.84[f]
Inhaled Aztreonam[c]	0.35	0.29	0.39	0.36[f]
Any Inhaled Antibiotics[c]	0.61	0.52	0.67	0.19[f]
Any Antibiotics[c]	0.82	0.74	0.86	0.17[f]
Ivacaftor[c]	0.044	0.071	0.028	0.36[g]
Ivacaftor-Lumacaftor[c]	0.018	0	0.028	0.53[g]
Ivacaftor or Ivacaftor-Lumacaftor[c]	0.061	0.071	0.056	0.71[g]

[a]Comparisons are between patients with no exacerbations and with one or more exacerbations during study year one.
[b]Median (range).
[c]Decimal fraction of patients.
[d]Linear regression.
[e]Quasipoisson regression.
[f]χ-square test.
[g]Fisher's Exact Test.

(Figure S3C) with no impact on model interpretations (Figure S3D), suggesting that cell counts likely do not mediate inflammation. The results also indicate that cell counts are not helpful when regarded as potential confounders and adjustment variables by themselves or in addition to age and prior exacerbations. Analyses replacing total cell counts with polymorphonuclear cell (**PMN**) counts produced similar results and identical interpretations. Thus total cell and PMN counts were excluded from further models.

S100A8, IL-10, IL-6, GMCSF, IFNγ, and ICAM1 were associated with weight-for-age Z score ($p < 0.05$) (Figure S4A), but only S100A8 retained significance indicated by an FDR < 0.001 after analysis (Figure S4B). Bivariable proportional hazards models of each biomarker with weight-for-age Z score for time to next exacerbation did not reduce the significance of individual biomarker relationships (Figure S4C) or substantively modify FDR analysis results prior to the evaluation of its effects in addition to adjustments for age and number of prior year exacerbations. The findings strongly suggest that weight-for-age Z score does not mediate inflammatory effects.

Further testing alongside adjustments by age and number of prior year exacerbations found that weight-for-age Z score may be an additional confounder of biomarker relationships (Figures S4C and S4D). Adding adjustment by weight-for-age Z score improved the false discovery analysis results for ENRAGE and MPO by placing them into the category of FDR < 0.05 and moved IL-17A into the category of FDR < 0.2; but effect sizes were not substantially changed. Several other biomarker results were slightly decreased in significance before

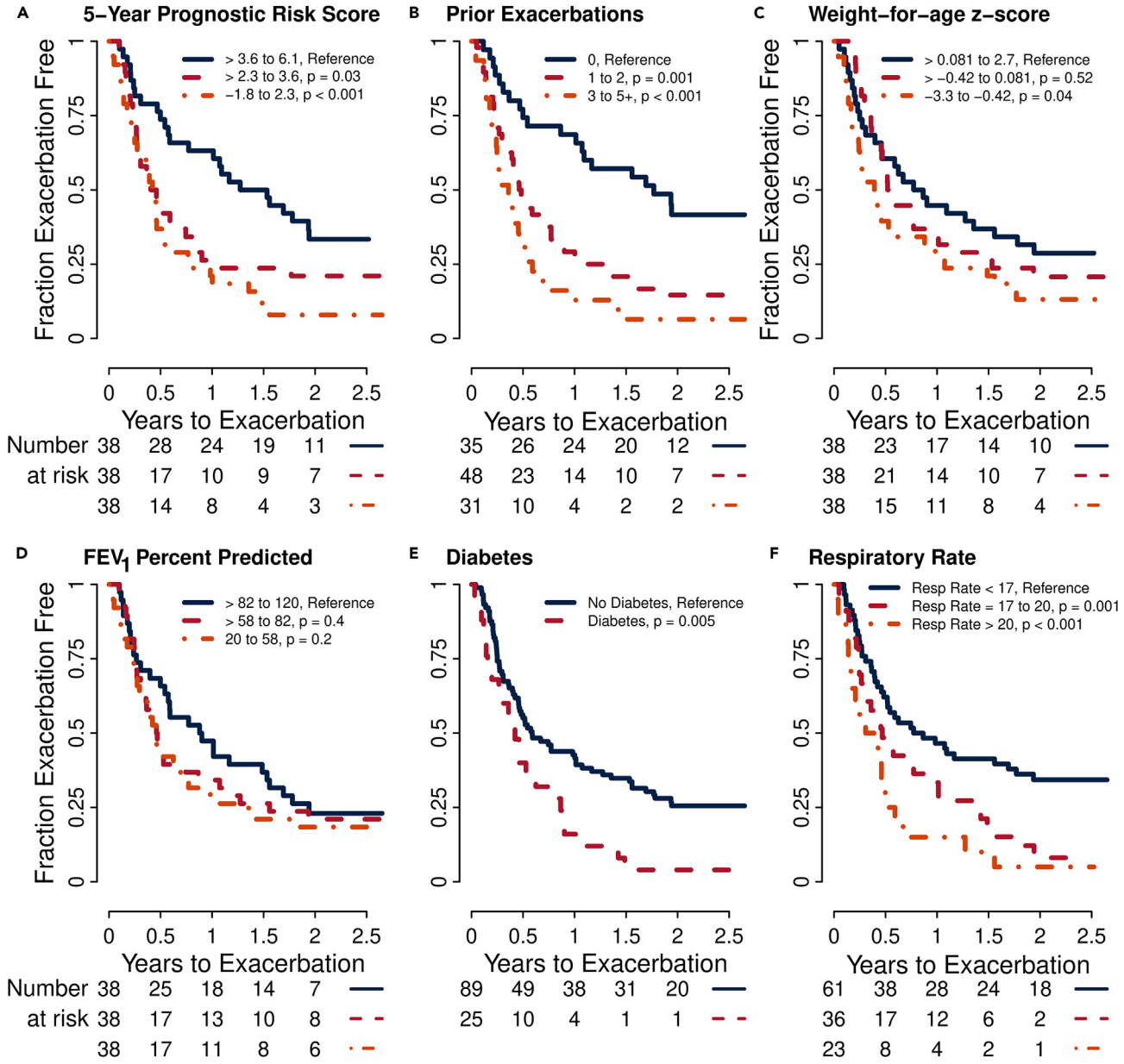


Figure 2. Kaplan-Meier plots exploring relationships of clinical factors with time to next pulmonary exacerbation

Patients were stratified into evenly sized groups when possible in each Number at Risk legend for (A) 5-year prognostic risk score.

(B) Number of pulmonary exacerbations in the year prior to enrollment.

(C) Weight-for-age Z score.

(D) FEV_1%.

(E) CF-related diabetes status and (F) respiratory rate. Measurements and status were from the day of enrollment. Patients were followed until occurrence of the next pulmonary exacerbation and censored. Groups are unequal in size for number of prior exacerbations, diabetes and respiratory rate because those values are ordinal and do not allow more even distribution. Results of log rank tests are shown in each panel legend. See also Table S2.

and after false discovery analysis without substantial changes in either effect sizes or their interpretations. A clear confounder effect for weight-for-age Z score is somewhat unexpected because of the paucity of significant associations with the biomarkers (Figure S4B), and the combination of expectation with variable small effects on biomarker effects and indicators of significance led us to exclude weight-for-age Z score from further models.

Diabetes and respiratory rate variables had few significant associations with inflammatory biomarkers (Figures S5 and S6). Bivariable proportional hazards models of individual biomarkers and diabetes found diminished significance for all biomarkers except ENRAGE. In false

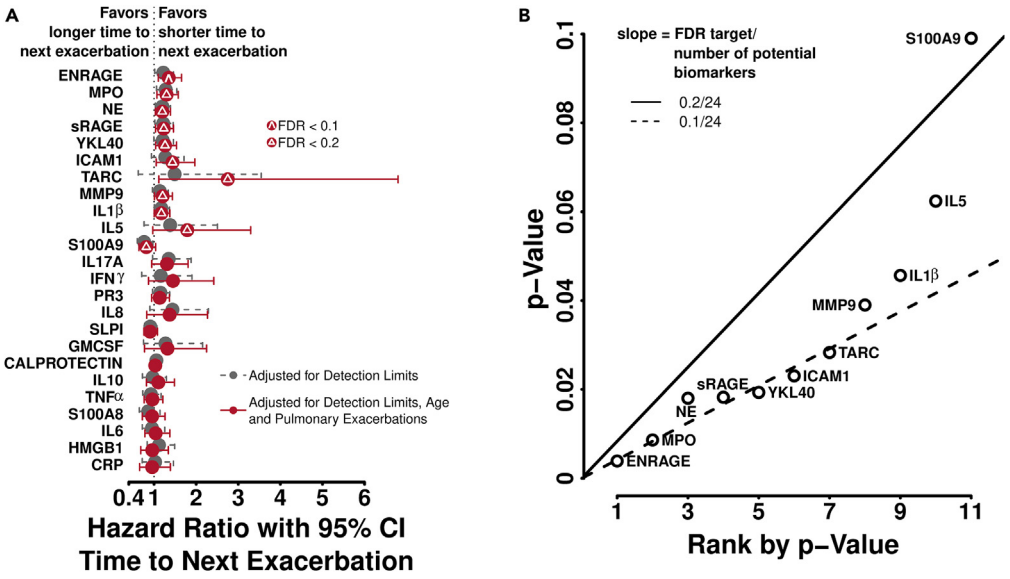


Figure 3. Proportional hazards models of time to next exacerbation

(A) Forest plots of hazard ratios with 95% confidence intervals derived from biomarker values as univariables adjusted for assay detection limits (gray) or assay detection limits and age and number of prior year pulmonary exacerbations as confounders (red). The upper 95% confidence limit for TARC adjusted by age and prior exacerbations is 6.80.

(B) Graphical FDR analysis was applied to hazard ratios adjusted by age and exacerbations. Using the p values for each hazard ratio, we ranked the potential biomarkers and drew lines with slopes determined by the thresholds for false discovery (set to 0.1 and 0.2) divided by the number of potential biomarkers in the entire study. A biomarker falling below a threshold line has a fractional chance of being a true finding, which is greater than *1—FDR threshold* for that line. Once a biomarker is plotted above a threshold line, no biomarkers with a larger rank are considered to be below that threshold. ENRAGE falls below the FDR 0.1 line, whereas MPO is just above, hence the next eight are considered below the FDR 0.2 threshold and not below FDR 0.1.

discovery analyses, only ENRAGE retained an FDR < 0.2. These findings suggest that diabetes potentially mediates inflammation and should be excluded from further models. Respiratory rate had no significant associations with any biomarker and had no further effect on models already adjusted by age and number of prior pulmonary exacerbations (Figure S6C). Multiple biomarkers were associated with heart rates (Figures S7A and S7B). Heart rate was a possible mediator of inflammation as inclusion in models of time to next exacerbation generally reduced biomarker effect sizes and significance (Figures S7C and S7D). These results prompted us to exclude CF-related diabetes, respiratory rate, and heart rate from further models.

Sensitivity analyses

Adjustments for inflammation modifying treatments did not substantially modify any biomarker hazard ratio. Use of corticosteroids, inhaled, oral, or both; chronic azithromycin; other oral antibiotics; and inhaled aztreonam and tobramycin, each alone or alternating every other month, had no significant associations with time to next exacerbation nor interactions with biomarker variables. Biomarker effects (Figure 3) remained stable throughout testing, suggesting that results are not due to these inflammation altering treatments.

Seven participants used ivacaftor or ivacaftor-lumacaftor, but usage of these medications had no independent effect or interaction with biomarker effects. FEV_1% recalculated with Global Lung Initiative equations[41] produced similar results and identical conclusions to using NHANES III.[42] Use of multiple strategies and alternative estimating equations for FEV_1%[43] for the participants belonging to groups for whom no specific equations exist had no impact on our findings.

We repeated all models to understand the impact of partially or completely missing biomarker information. Out of a possible 2,736 total biomarker measurements, there were 25 values below the lower limit of detection (0.9%), 33 above the upper limit of detection (1.2%), and 51 completely missing (1.8%). The exact biomarker values partially or completely missing are enumerated in Table S6. We used biomarker datasets that either omitted records with missingness or imputed the missing values. Values were similar across the original and alternative datasets (Table S6). Results of all models using datasets with records omitted for missingness or with imputed values were similar to the original models (Tables S7–S13), and no interpretations changed with use of any of the alternative datasets created to address missingness.

DISCUSSION

We measured 24 sputum biomarkers from 114 randomly chosen, clinically stable adolescents and adults to prospectively validate potentially causal inflammatory biomarker relationships with future pulmonary exacerbations in a cohort representative of people with CF in the US from the study era.[33] The complexity of disease required careful explorations of confounding[30] and mediation[31] to improve causal inference

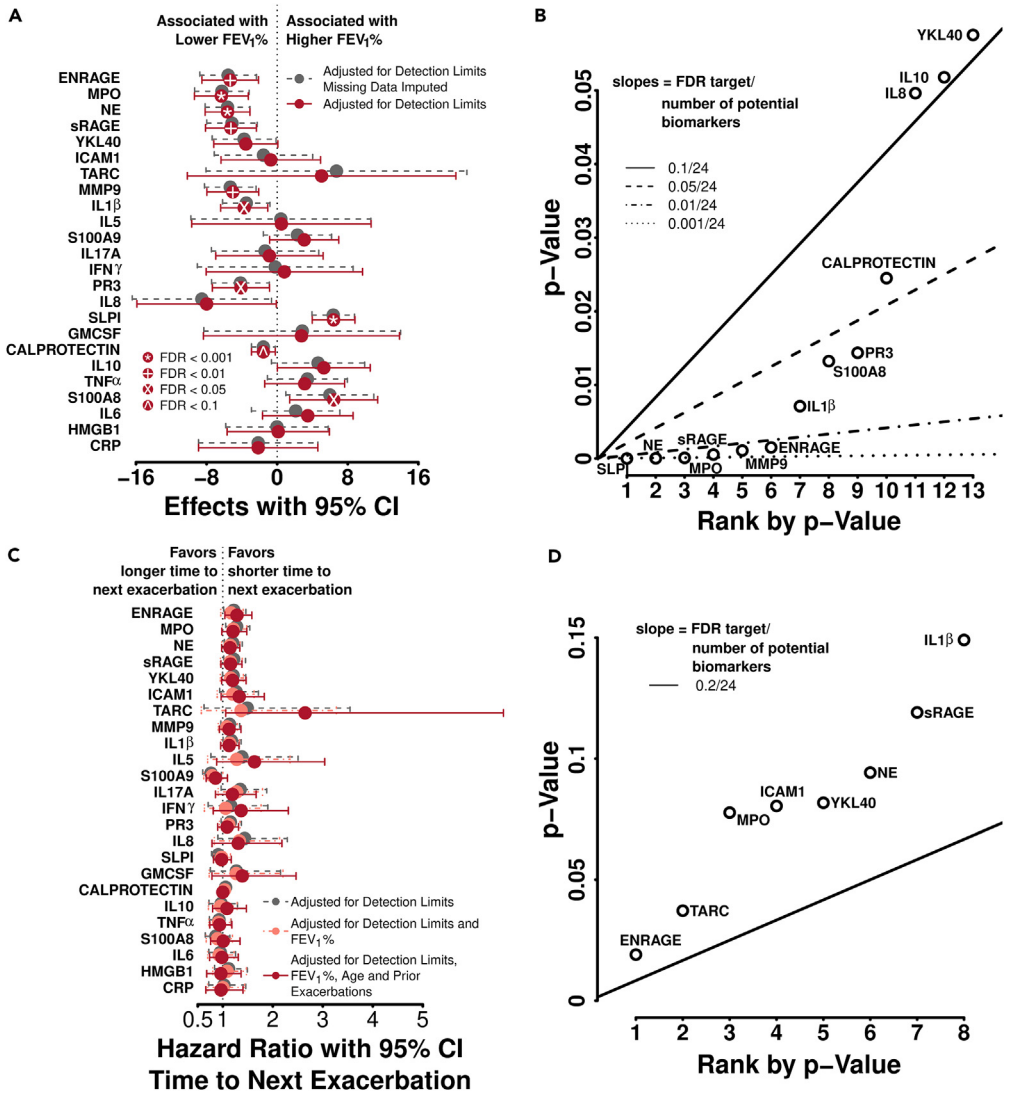


Figure 4. Relationships between biomarkers and FEV_1%

(A) Forest plots show linear regression model effect estimates and 95% confidence intervals for associations between inflammatory biomarkers adjusted for assay detection limits and FEV_1%.

(B) FDR analysis shows nine biomarkers retain significance for associations with FEV_1%, three each at FDR < 0.001, FDR < 0.01, and FDR < 0.05.

(C) Significance of hazard ratios of biomarkers when used in models with FEV_1% for time to next pulmonary exacerbation (**gray**) is reduced during testing and confirmation of FEV_1% as a mediator (**salmon**) even when adjusted by age and prior pulmonary exacerbations (**red**). (See Figure S2 for the results of FDR analysis of unadjusted bivariable models.) The order of biomarkers allows easier comparison with Figure 2A.

(D) In mediation testing of FEV_1%, no biomarkers used as adjustments to FEV_1% for proportional hazards models of time to next pulmonary exacerbation clearly retain statistical significance after FDR analysis with FDR < 0.2.

quality.[37,44] Among the three potential biomarkers previously studied specifically for relationships with future exacerbations,[17–19] NE was most strongly related in validation, whereas there were surprisingly weak relationships for HMGB1 and calprotectin. However, we found additional strong positive associations with time to next exacerbation for biomarkers that were previously found to have associations with past or concurrent exacerbations (namely ENRAGE, MPO, sRAGE, ICAM1, YKL40, TARC, MMP9, IL-1β, and IL-5).[20–28] Altogether, these biomarkers suggest three types of inflammation, RAGE axis, protease-mediated, and oxidative stress, that may be attractive treatment targets, perhaps in combination, for anti-inflammatory bench studies and clinical trials to help prevent pulmonary exacerbations and treat underlying inflammation.

We found that SLPI, NE, and MPO (Figures 4A and 4B) were associated with FEV_1% with an FDR < 0.001. With FDR < 0.01, we found associations with sRAGE, MMP9, and ENRAGE. SLPI was associated with higher, whereas NE, sRAGE, MMP9, and ENRAGE were associated with lower, FEV_1% and shorter times to next exacerbation (Figure 3A).

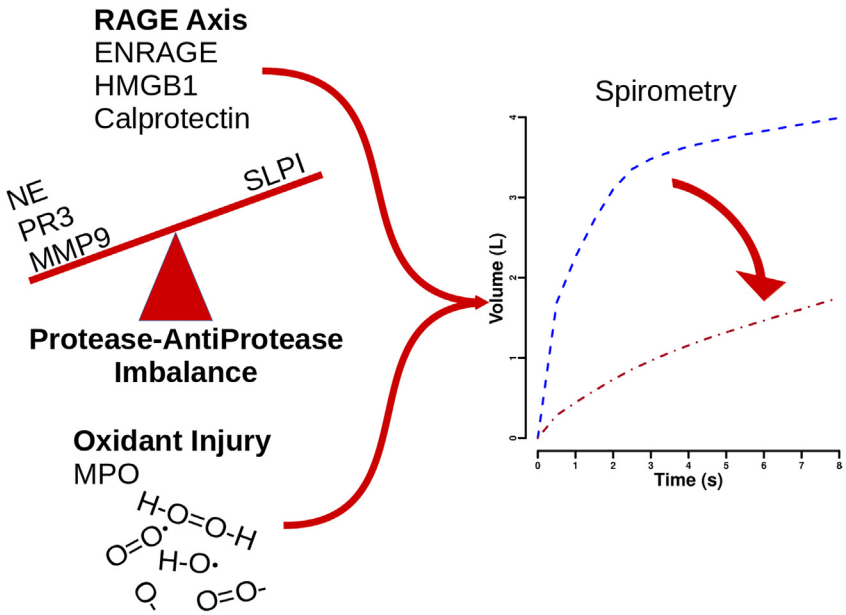


Figure 5. Relationship between RAGE axis, protease-antiprotease imbalance, oxidant injury, and lung function

Three pathways of injury with representative biomarkers shown contribute to reducing lung function. Sharp decreases in lung function are the most frequent indicator used by clinicians making a diagnosis of pulmonary exacerbation of CF.

SLPI, NE, and MMP9 represent opposing sides in protease-antiprotease-imbalance-associated inflammation. SLPI inhibits NE and other proteases such as cathepsin G and suppresses MMP9 expression and other NF-κB mediated pro-inflammatory activities,[18,45,46] although it does not inhibit proteinase 3 (**PR3**).[47] In opposition,[48] NE released from neutrophil azurophilic granules degrades pericellular matrix in a manner that is impossible to completely stop.[49,50] NE is a major cause of inflammation in multiple lung diseases[51] and specifically in pulmonary exacerbations in CF.[18,52]

MPO was associated with lower FEV_1% and shorter time to next exacerbation confirming prior findings.[53] Upon neutrophil azurophilic granule release, it generates antibacterial reactive oxygen species and hypohalous acids.[54] However, the reactive oxygen species can cause cellular and airway oxidative damage.[46] MPO is representative of oxidative stress resulting from inflammation.

ENRAGE and sRAGE were associated with decreased FEV_1% and time to next exacerbation. Both are receptors for advanced glycation end products (**RAGE**) involved in RAGE-axis-related inflammation.[21,55] These findings suggest that RAGE-axis-, protease-anti-protease-imbalance-, and reactive-oxygen-species-related inflammation are related to lower FEV_1% (Figure 5) and shorter times to next exacerbation in CF.

Derivation of relationships between biomarkers and time to next exacerbation required evaluation of confounding and mediating relationships involving clinical variables. Confounders are associated with both biomarkers and outcomes (Figure 1) but are not part of causal disease pathways. However, at least some are needed to adjust models for non-causal relationships.[30,31] We adjusted our proportional hazards models by including age and number of prior pulmonary exacerbations. Weight for age Z score and sex had small or no confounding effects and no impact on our interpretations of biomarker relationships with time to next exacerbation. We excluded these and other low-impact confounders in favor of increased parsimony and interpretability of our biomarker models.

Mediators are intermediate outcomes between putative causal factors and targeted clinical outcomes such as time to next exacerbation. Mediators may obscure upstream causal pathway effects,[37,44] frustrating identification of potentially treatable pathologic relationships. Exclusion of mediators from inference models improves evaluation of those relationships and primes studies seeking more comprehensive treatments[30,31] (Figure 1). For example, in a study of inflammation responses to intravenous antibiotic treatment of lower airway infections in 19 subjects with CF, protease-antiprotease balance returned toward normal;[26] however, in a similar study, other neutrophil-related inflammation types persisted,[25] suggesting that concurrent treatments for multiple forms of inflammation may improve outcomes.

Statistical inference based on probabilistic models is rarely convincing for causality (cigarette smoking causing lung cancer[56] is an important exception) but can focus study on targets likely to be causal.[37] This study reproduced pre-study expectations[17,19] of strong associations between calprotectin and HMGB1 with prior pulmonary exacerbations (Table S5); however, they were not associated with time to next pulmonary exacerbation. Inability to validate as potentially causal suggests that clinical trials antagonizing calprotectin and HMGB1 to prevent pulmonary exacerbations will fail.

However, after FDR analysis,[39] our testing found 10 biomarkers that may be involved in causal pathways of inflammation and explanatory for pulmonary exacerbations in CF (Figure 3). Setting FDR < 0.2 suggests that eight of the ten are likely associated with time to next pulmonary exacerbation (Figure 3). The findings validate NE as a potentially causal contributor to exacerbations. NE was previously associated with future

exacerbations,[18] and our results suggest protease-anti-protease imbalance as an underlying mechanism mediated through an effect on FEV_1%.

The other biomarkers associated with time to next exacerbation in our study were all previously associated in some way with prior or concurrent exacerbations, and they fall into several categories. ENRAGE and sRAGE are constituents of RAGE-axis-related inflammation. ENRAGE is a high-avidity RAGE ligand[57] elevated in CF airways in infancy[58] associated with low FEV_1 and CF-related diabetes.[55] It increases with pulmonary exacerbation and falls with antibiotic treatment.[21] Beyond CF, ENRAGE is associated with cigarette smoking,[59] coronary artery[60] and fatal kidney[61] diseases, and development and resolution of acute respiratory distress syndrome (**ARDS**).[62]

ENRAGE was strongly correlated with MMP9 and MPO (Table S3), suggesting networked expression of RAGE axis, protease imbalance, and oxidative stress inflammation. Correlations between sRAGE and NE (0.80), MMP9 (0.65), and IL-1β (0.69) strengthen possible associations between RAGE and protease imbalance (Table S3). MPO, ICAM1, NE, YKL40, TARC, and MMP9 reflect the participation of PMN,[25,27,28] PMN-associated proteases,[18,63] and reactive oxygen species[46] in CF airway inflammation (Figure 5).

Severe CF lung disease is steadily decreasing as CFTR modulator usage increases among those with eligible mutations.[64] FEV_1% remains the single most common measure of disease severity but poorly distinguishes between rapid or slow worsening of lung disease,[29] and it mediates and obscures biomarker effects (Figure 4; Table S8). As lung disease severity lessens in the CF population,[65] FEV_1% will become less useful as an investigational tool. In contrast, potentially causal biomarkers (as well as strongly associated but probably non-causal biomarkers such as calprotectin) may enable more precise tracking of disease and interventional trial outcomes. Focusing on potentially causal biomarkers will amplify the clinical relevance of bench research and the effectiveness, timeliness, and precision of animal studies and interventional human trials.

Extrapulmonary manifestations of CF are increasing as individuals with CF reach older ages in larger numbers[66,67]; some of these manifestations may modify or reflect airway inflammation during a pulmonary exacerbation. Even accounting for efforts to improve detection, diabetes is increasing in frequency at all ages in the CF population.[14] People with CF who also have CFRD are twice as likely to suffer exacerbations and generally have worsened lung disease,[68] prior findings that our results support despite a much smaller dataset; 22 of the 114 participants with CF also had CFRD (19%), which was associated with earlier exacerbation of lung disease (Figure 2E). Nevertheless, we found no direct connections between the biomarkers we studied and CFRD and little direct effect on biomarker relationships with time to next exacerbation (Figure S5). This set of findings suggests that additional factors relating systemic diabetes and exacerbations exist but remain undiscovered.

Many of the biomarkers that we measured from sputum are found in blood as reporters of systemic inflammation in the setting of pulmonary exacerbations. Inflammatory oxidative stress and other inflammatory markers are elevated during pulmonary exacerbations in blood and plasma and respond rapidly to treatment.[25,69] However, airway inflammation persists when systemic inflammation subsides after clinical resolution of an exacerbation,[25] providing information on recent exacerbations that may contribute to and explain the success of our model adjustments with prior year exacerbation counts. Inflammatory markers such as assessments of up- and downregulation of gene expressions related to lung inflammation are more varied in airway PMN compared with blood circulating PMN,[70] increasing the information available from airway inflammatory biomarker measurements compared with blood measurements.[71] In our study, we collected biomarker-containing sputum in order to minimize any reluctance to participate due to blood draw associated pain. Based on prior data, we likely improved our ability to detect relationships between inflammatory biomarkers during clinical stability compared with use of systemic markers from blood.

Limitations of the study

Sputum is a non-invasive but complex sample that contains multiple inflammation-related components. We focused in this work on aqueous human proteins that we carefully isolated and measured, but the list of proteins and human-derived molecules is incomplete; some excluded molecules may be important for inclusion in further studies.[72] We excluded pellet fractions from our samples that contain microbes, microbial byproducts, human cell fragments, and other debris, and we excluded the lipid aliquots from the samples that contain extracellular vesicles that may help maintain homeostasis or contain inflammation-related proteins, micro RNA, and other constituents such as vesicle surface-bound neutrophil elastase that remains active but becomes resistant to inhibition by anti-proteases.[73–76]

Because we pursue explanatory models that may reveal new targets for treatment,[77] no methods of model selection are sufficient[30] to confidently produce a multivariable panel of inflammatory biomarkers to better understand the evolution of increasing inflammation and future pulmonary exacerbations. Multivariable model selection using forward, backward, or other input variable selection methods with or without automated procedures may be adequate for developing prediction models that need to maximize accuracy above all else for prognostication, but these methods do not necessarily produce models that provide insights into underlying mechanisms of disease.

An effort to derive an explanatory model including multiple inflammatory biomarkers would need to consider that some biomarkers may confound while others mediate the effects of a given set of biomarkers. An explanatory model would require characterization of potential webs of confounding and mediation among the biomarkers prior to selection of final multi-variable models. The time required even for fully automated procedures involving 24 biomarkers is not feasible. Thus, for an exploratory investigation, we considered each biomarker in turn in univariable models and in models adjusted by clinical confounders and excluding mediators. This work exploring individual biomarkers may provide useful guidance prior to launching new observational, bench, and interventional investigations, perhaps with factorial designs,[78] to prospectively test our individual biomarkers for potential causal relationships with future increasing airway inflammation and clinical pulmonary exacerbations.

Since the study was initiated, highly effective CFTR modulator therapies beyond ivacaftor alone or combination ivacaftor-lumacaftor have been broadly adopted among genetically eligible patients. These agents may reduce inflammation, associated infections, and sputum production for biomarker sampling. Early data, however, suggest that inflammation and infections persist in part even in these patients,[7–10] and remaining sputum production may suffice for serial study.[79] Although many of our study patients took every available effective CF treatment, due to study timing, only 6% received CFTR modulators. In this study, we found a non-significant decrease of approximately 50% on average in sputum weight collected among participants using either ivacaftor or ivacaftor-lumacaftor treatments (Figure S8). Adjustments for lung function or prior exacerbations had no impact on this finding. Two out of thirteen missing HMGB1 measurements were from samples that were from participants treated with modulators but none of the other missing biomarker measurements were from modulator-treated individuals. Our results were statistically insensitive to CFTR modulator use, but our findings do not settle questions whether usage substantially reduces airway inflammation, reduces measurability of any specific inflammatory marker, or prevents collection of sputum for study. However, among patients who are genetically ineligible or otherwise unable to obtain modulator therapy, there is no doubt that inflammation remains an important investigation and treatment target.

Finally, our explanatory models infer but cannot confirm causal relationships[37,44] arising from an inflammatory biomarker network. RAGE-axis inflammation, protease-mediated injury, and oxidative stress are plausibly associated with future exacerbation. FEV_1%, the single most important CF survival predictor,[14,29] statistically mediates the association of inflammatory markers with time to next pulmonary exacerbation, but a role as a statistical mediator may indicate that FEV_1% is a summarizing reporter of biological inflammation of different kinds in actual airways. Further studies in cell culture, animal models, or human trials of anti-proteases, anti-oxidants, RAGE inhibiting agents, or a combination are needed to confirm our inferences and potentially develop new treatments for CF airway inflammation. Because these inflammatory processes are similar and important in other lung diseases such as bronchiectasis and ARDS, progress in CF with novel anti-inflammatory treatments may have broad application.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Ethical reviews
  - Study participants
- METHOD DETAILS
  - Pre-study preparations and personnel training
  - Study design, participants and sample size
  - Sputum and biomarker collections and measurements
  - ELISA procedures for measurements of calprotectin, CRP and HMGB1
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical software
  - Statistical analysis procedures
  - Assessments of confounding and mediation
  - Development of explanatory models and false discovery rate analyses
  - Sensitivity analyses
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.108835.

ACKNOWLEDGMENTS

We remember the warmth, generosity, and skill of our colleagues who died during the course of the work reported here: Dr Barbara Allison Chatfield, August 25, 1959 – October 8, 2016 and Professor Sir David Roxbee Cox, July 15, 1924 – January 18, 2022. We miss them both greatly.

We thank the study participants and their families. We thank Mike Mitchell in the Pulmonary Laboratory at the University of Utah for technical assistance.

This project was supported by the CF Foundation (CFF) (LIOU13A0, LIOU14Y4), the National Center for Advancing Translational Sciences at the National Institutes of Health (NCATS/NIH 8UL1TR000105 [formerly UL1RR025764]), the Ben B and Iris M Margolis Family Foundation of Utah, and the Claudia Ruth Goodrich Stevens Endowment Fund.

AUTHOR CONTRIBUTIONS

T.G.L. is the manuscript guarantor and takes responsibility for the integrity of the work from inception to publication T.G.L. drafted the initial manuscript; all authors contributed interpretive insights, edits, and corrections. M.H. created a graphical abstract with assistance from T.G.L. and assisted T.G.L., R.D.S., and F.R.A. to diagram interactions between biomarkers and lung disease. D.R.C. authored and designed many of the methods cited and used in this work and provided extensive guidance drawn from decades of experience on statistical causal inference and study design. He participated in revisions of the work during the last days and weeks of his life. D.R.C. initiated and oversaw development of the statistical plan for randomized enrollment by T.G.L., F.R.A., and R.K. with additional inputs from J.L.J., K.R.P., and S.D.S. T.G.L., F.R.A., D.R.C., C.K., R.K., and R.D.S. performed the statistical analysis. J.L.J. oversaw IRB submissions throughout the MWCFC with help at each site from N.A., D.D., J.A.F., B.G., C.M.K., O.M., K.A.P., K.R.P., A.J.R., J.B.V., and R.Y. T.G.L., N.A., P.S.B., B.A.C., C.L.D., D.D., J.A.F., B.G., J.L.J., T.H., J.R.H., C.M.K., Y.L., J.L., N.L., O.M., C.N., K.A.P., K.R.P., A.L.Q., P.R., A.J.R., S.D.S., J.L.T.-C., J.B.V., R.Y., J.P.C., J.S.E., and K.N.O. developed and performed patient enrollment processes, sample collections, and clinical annotations. J.L.J. oversaw the work preceding the study planning meeting with critical assistance from K.A.P., J.A.F., and J.B.V. J.L.J. oversaw and coordinated logistics and regulatory documentation for the entire study. T.G.L., J.A.F., J.L.J., Y.L., K.A.P., and J.B.V. developed the sputum processing protocol. J.A.F., J.L.J., K.A.P., Y.L., and J.B.V. provided central and on-site laboratory training. T.G.L. and J.L.J. performed site initiation visits. J.L.J. and K.A.P. centrally monitored data collection. J.A.F., K.A.P., and J.B.V. performed interim and end-of-study site visits to verify data integrity and regulatory compliance. K.A.P. developed and implemented REDCap electronic data capture[80] with input from T.G.L., J.A.F., J.L.J., and J.B.V. K.A.P. oversaw data entry, query generation and resolution, and generation of raw data reports. J.B.V. with assistance from J.A.F., J.L.J., and Y.L. developed protocols and oversaw specimen shipping by N.A., D.D., B.G., C.M.K., O.M., K.R.P., A.J.R., and R.Y. J.A.F., J.L.J., J.B.V., K.A.P., and Y.L. received, handled, and successfully managed sputum specimens centrally despite shipment damage, floods, power outages, freezer, and complete building services failures. Y.L. performed ELISA studies in Utah. S.S. oversaw NE activity determinations in Colorado. T.G.L., F.R.A., and J.L.T.-C. conferred on the use of FEV_1% estimating equations to ensure inclusion of all participants in all analyses. T.G.L. obtained funding with assistance from F.R.A., J.L.J., and S.D.S. and additional help from P.S.B., B.A.C., C.L.D., R.K., J.L., C.N., and D.R.C. J.P.C., J.S.E., and K.N.O. were the Study Advisory Committee and contributed to multiple study aspects. All authors had access to complete raw and processed datasets from this study.

The results, interpretations, and opinions expressed in this work are those of the authors and do not necessarily represent the views of their institutions, the NIH, NHLBI, NCATS, NSF, the Veteran's Administration, 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..

DECLARATION OF INTERESTS

T.G.L., J.A.F., J.L.J., Y.L., K.A.P., and J.B.V. received other support from the CFF (CC132-16AD, 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, Calithera Biosciences, Corbus Pharmaceuticals, Gilead Sciences, Laurent Pharmaceuticals, Nivalis Therapeutics, Novartis, Proteostasis, Savara Pharmaceuticals, Translate Bio, and Vertex Pharmaceuticals. F.R.A. received additional other support from the NHLBI/NIH (R01 HL125520), the National Science Foundation (EMSW21-RTG), and the Margolis Family Foundation of Utah. P.S.B. 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 Therapeutics. B.A.C. received other support from the CFF (C112–12, C112-TDC09Y, 10063SUB, 41339154.s132P010379SUB) and received support for clinical trials from Genentech, Novartis, and Vertex Pharmaceuticals. C.L.D. received other support from the CFF (C004–11, C004-TDC09Y, DAINES11Y3) and from the Health Resources and Services Administration (T72MC00012). J.A.F. transitioned to become an employee of ICON plc, a clinical research organization involved in various trials pertinent to CF during the study and is now an employee of Vertex Pharmaceuticals; JAF, ICON, and Vertex had no direct involvement in performance of the study following the initial change in affiliation. T.H. received other support from the CFF (PACE, Center Grant) and received support for clinical trials from Celtaxsys and Vertex Pharmaceuticals. J.R.H. received other support from the NHLBI/NIH (HHSN268200900018C) and the Veterans Administration Healthcare System (I01 BX001533). J.L. received other support from the CFF (C017-11AF). C.N. received other support from the CFF (C138-12). P.R. received other support from the CFF (C003–12, C003-TDC09Y). S.D.S. received other support from the CFF (AQUADEK12K1, SAGEL11CS0, GOAL13K2, NICK13A0, SAGEL14K1, NICK15R0) and the NHLBI/NIH (U54 HL096458) and the NCATS/NIH (Colorado CTSA Grant Number UL1 TR002535). J.L.T.-C. received other support from the CFF (TDC) and the NHLBI/NIH (HL103801) and received support for clinical trials from Vertex Pharmaceuticals. K.N.O. was funded by the intramural research program of the NHLBI, NIH, during the study. Neither the project sponsors nor any sources of other support had direct roles in development and conduct of the study. None of the sponsors of clinical trials mentioned earlier participated in any way with this trial.

Received: September 18, 2023
Revised: November 2, 2023
Accepted: January 3, 2024
Published: January 9, 2024

REFERENCES

1. Zemanick, E.T., Wagner, B.D., Robertson, C.E., Ahrens, R.C., Chmiel, J.F., Clancy, J.P., Gibson, R.L., Harris, W.T., Kurland, G., Laguna, T.A., et al. (2017). Airway microbiota across age and disease spectrum in cystic fibrosis. Eur. Respir. J. *50*, 1700832. https://doi.org/10.1183/13993003.00832-2017.
2. Granchelli, A.M., Adler, F.R., Keogh, R.H., Kartsonaki, C., Cox, D.R., and Liou, T.G. (2018). Microbial Interactions in the Cystic Fibrosis Airway. J. Clin. Microbiol. *56*, e00354-18. https://doi.org/10.1128/JCM.00354-18.
3. Flume, P.A., Wainwright, C.E., Elizabeth Tullis, D., Rodriguez, S., Niknian, M., Higgins, M., Davies, J.C., and Wagener, J.S. (2018). Recovery of lung function following a pulmonary exacerbation in patients with cystic fibrosis and the G551D-CFTR mutation treated with ivacaftor. J. Cyst. Fibros. *17*, 83–88. https://doi.org/10.1016/j.jcf.2017.06.002.
4. Tunney, M.M., and Wark, P. (2023). Long-term therapy with elexacaftor/tezacaftor/ivacaftor (ETI) in cystic fibrosis: improved clinical outcomes but infection and inflammation persist. Eur. Respir. J. *62*, 2301008. https://doi.org/10.1183/13993003.01008-2023.
5. Casey, M., Gabillard-Lefort, C., McElvaney, O.F., McElvaney, O.J., Carroll, T., Heeney, R.C., Gunaratnam, C., Reeves, E.P., Murphy, M.P., and McElvaney, N.G. (2023). Effect of elexacaftor/tezacaftor/ivacaftor on airway and systemic inflammation in cystic fibrosis. Thorax *78*, 835–839. https://doi.org/10.1136/thorax-2022-219943.
6. Schaupp, L., Addante, A., Völler, M., Fentker, K., Kuppe, A., Bardua, M., Duerr, J., Piehler, L., Röhmel, J., Thee, S., et al. (2023). Longitudinal effects of elexacaftor/tezacaftor/ivacaftor on sputum viscoelastic properties, airway infection and inflammation in patients with cystic fibrosis. Eur. Respir. J. *62*, 2202153. https://doi.org/10.1183/13993003.02153-2022.
7. Lepissier, A., Bonnel, A.S., Wizla, N., Weiss, L., Mittaine, M., Bessaci, K., Kerem, E., Houdouin, V., Reix, P., Marguet, C., et al. (2023). Moving the Dial on Airway Inflammation in Response to Trikafta in Adolescents with Cystic Fibrosis. Am. J. Respir. Crit. Care Med. *207*, 792–795. https://doi.org/10.1164/rccm.202210-1938LE.
8. Caverly, L.J., Riquelme, S.A., and Hisert, K.B. (2022). The Impact of Highly Effective Modulator Therapy on Cystic Fibrosis Microbiology and Inflammation. Clin. Chest Med. *43*, 647–665. https://doi.org/10.1016/j.ccm.2022.06.007.
9. Gillan, J.L., Davidson, D.J., and Gray, R.D. (2021). Targeting cystic fibrosis inflammation in the age of CFTR modulators: focus on macrophages. Eur. Respir. J. *57*, 2003502. https://doi.org/10.1183/13993003.03502-2020.
10. Harris, J.K., Wagner, B.D., Zemanick, E.T., Robertson, C.E., Stevens, M.J., Heltshe, S.L., Rowe, S.M., and Sagel, S.D. (2020). Changes in Airway Microbiome and Inflammation with Ivacaftor Treatment in Patients with Cystic Fibrosis and the G551D Mutation. Ann. Am. Thorac. Soc. *17*, 212–220. https://doi.org/10.1513/AnnalsATS.201907-493OC.
11. Dagenais, R.V.E., Su, V.C.H., and Quon, B.S. (2020). Real-World Safety of CFTR Modulators in the Treatment of Cystic Fibrosis: A Systematic Review. J. Clin. Med. *10*, 23. https://doi.org/10.3390/jcm10010023.
12. Lopes-Pacheco, M. (2019). CFTR Modulators: The Changing Face of Cystic Fibrosis in the Era of Precision Medicine. Front. Pharmacol. *10*, 1662. https://doi.org/10.3389/fphar.2019.01662.
13. Elborn, J.S. (2016). Cystic fibrosis. Lancet *388*, 2519–2531. https://doi.org/10.1016/S0140-6736(16)00576-6.
14. Liou, T.G., Kartsonaki, C., Keogh, R.H., and Adler, F.R. (2020). Evaluation of a five-year predicted survival model for cystic fibrosis in later time periods. Sci. Rep. *10*, 6602. https://doi.org/10.1038/s41598-020-63590-8.
15. Cystic Fibrosis Foundation Patient Registry 2021 Annual Data Report (2022 (Cystic Fibrosis Foundation).
16. Bell, S.C., Mall, M.A., Gutierrez, H., Macek, M., Madge, S., Davies, J.C., Burgel, P.-R., Tullis, E., Castaños, C., Castellani, C., et al. (2019). The future of cystic fibrosis care: a global perspective. Lancet Respir. Med. *8*, 65–124. https://doi.org/10.1016/S2213-2600(19)30337-6.
17. Liou, T.G., Adler, F.R., Keogh, R.H., Li, Y., Jensen, J.L., Walsh, W., Packer, K., Clark, T., Carveth, H., Chen, J., et al. (2012). Sputum biomarkers and the prediction of clinical outcomes in patients with cystic fibrosis. PLoS One *7*, e42748. https://doi.org/10.1371/journal.pone.0042748.
18. Sagel, S.D., Wagner, B.D., Anthony, M.M., Emmett, P., and Zemanick, E.T. (2012). Sputum biomarkers of inflammation and lung function decline in children with cystic fibrosis. Am. J. Respir. Crit. Care Med. *186*, 857–865. https://doi.org/10.1164/rccm.201203-0507OC.
19. Reid, P.A., McAllister, D.A., Boyd, A.C., Innes, J.A., Porteous, D., Greening, A.P., and Gray, R.D. (2015). Measurement of Serum Calprotectin in Stable Patients Predicts Exacerbation and Lung Function Decline in Cystic Fibrosis. Am. J. Respir. Crit. Care Med. *191*, 233–236. https://doi.org/10.1164/rccm.201407-1365LE.
20. Lorenz, E., Muhlebach, M.S., Tessier, P.A., Alexis, N.E., Duncan Hite, R., Seeds, M.C., Peden, D.B., and Meredith, W. (2008). Different expression ratio of S100A8/A9 and S100A12 in acute and chronic lung diseases. Respir. Med. *102*, 567–573. https://doi.org/10.1016/j.rmed.2007.11.011.
21. Foell, D., Seeliger, S., Vogl, T., Koch, H.-G., Maschek, H., Harms, E., Sorg, C., and Roth, J. (2003). Expression of S100A12 (EN-RAGE) in cystic fibrosis. Thorax *58*, 613–617. https://doi.org/10.1136/thorax.58.7.613.
22. Mulrennan, S., Baltic, S., Aggarwal, S., Wood, J., Miranda, A., Frost, F., Kaye, J., and Thompson, P.J. (2015). The role of receptor for advanced glycation end products in airway inflammation in CF and CF related diabetes. Sci. Rep. *5*, 8931. https://doi.org/10.1038/srep08931.
23. Devereux, G., Steele, S., Jagelman, T., Fielding, S., Muirhead, R., Brady, J., Grierson, C., Brooker, R., Winter, J., Fardon, T., et al. (2014). An observational study of matrix metalloproteinase (MMP)-9 in cystic fibrosis. J. Cyst. Fibros. *13*, 557–563. https://doi.org/10.1016/j.jcf.2014.01.010.
24. Fischer, N., Hentschel, J., Markert, U.R., Keller, P.M., Pletz, M.W., and Mainz, J.G. (2014). Non-invasive assessment of upper and lower airway infection and inflammation in CF patients. Pediatr. Pulmonol. *49*, 1065–1075. https://doi.org/10.1002/ppul.22982.
25. Downey, D.G., Brockbank, S., Martin, S.L., Ennis, M., and Elborn, J.S. (2007). The effect of treatment of cystic fibrosis pulmonary exacerbations on airways and systemic inflammation. Pediatr. Pulmonol. *42*, 729–735. https://doi.org/10.1002/ppul.20646.
26. Müller, U., Hentschel, J., Janhsen, W.K., Hünniger, K., Hipler, U.-C., Sonnemann, J., Pfister, W., Böer, K., Lehmann, T., and Mainz, J.G. (2015). Changes of Proteases, Antiproteases, and Pathogens in Cystic Fibrosis Patients' Upper and Lower Airways after IV-Antibiotic Therapy. Mediat. Inflamm. *2015*, 626530. https://doi.org/10.1155/2015/626530.
27. Hartl, D., Griese, M., Kappler, M., Zissel, G., Reinhardt, D., Rebhan, C., Schendel, D.J., and Krauss-Etschmann, S. (2006). Pulmonary T(H)2 response in Pseudomonas aeruginosa-infected patients with cystic fibrosis. J. Allergy Clin. Immunol. *117*, 204–211. https://doi.org/10.1016/j.jaci.2005.09.023.
28. Fantino, E., Gangell, C.L., Hartl, D., and Sly, P.D.; AREST CF (2014). Airway, but not serum or urinary, levels of YKL-40 reflect inflammation in early cystic fibrosis lung disease. BMC Pulm. Med. *14*, 28. https://doi.org/10.1186/1471-2466-14-28.
29. Liou, T.G., Adler, F.R., Fitzsimmons, S.C., Cahill, B.C., Hibbs, J.R., and Marshall, B.C. (2001). Predictive 5-year survivorship model of cystic fibrosis. Am. J. Epidemiol. *153*, 345–352. https://doi.org/10.1093/aje/153.4.345.
30. Lederer, D.J., Bell, S.C., Branson, R.D., Chalmers, J.D., Marshall, R., Maslove, D.M., Ost, D.E., Punjabi, N.M., Schatz, M., Smyth, A.R., et al. (2019). Control of Confounding and Reporting of Results in Causal Inference Studies. Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals. Ann. Am. Thorac. Soc. *16*, 22–28. https://doi.org/10.1513/AnnalsATS.201808-564PS.
31. Baron, R.M., and Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J. Pers. Soc. Psychol. *51*, 1173–1182. https://doi.org/10.1037//0022-3514.51.6.1173.
32. Jung, S.J. (2021). Introduction to Mediation Analysis and Examples of Its Application to Real-world Data. J Prev Med Public Health *54*, 166–172. https://doi.org/10.3961/jpmph.21.069.
33. Liou, T.G., Adler, F.R., Argel, N., Asfour, F., Brown, P.S., Chatfield, B.A., Daines, C.L., Durham, D., Francis, J.A., Glover, B., et al. (2019). Prospective multicenter randomized patient recruitment and sample collection to enable future measurements of sputum biomarkers of inflammation in an observational study of cystic fibrosis. BMC Med. Res. Methodol. *19*, 88. https://doi.org/10.1186/s12874-019-0705-0.
34. Kaplan, E.L., and Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. J. Am. Stat. Assoc. *53*, 457–481. https://doi.org/10.2307/2281868.
35. Cox, D.R. (1958). The Regression Analysis of Binary Sequences. J. Roy. Stat. Soc. B *20*, 215–232.
36. Cox, D.R. (1972). Regression models and life-tables. J R Stat Soc B Met *34*, 187–202. https://doi.org/10.2307/2985181.
37. Cox, D.R., and Wermuth, N. (2004). Causality: A statistical view. Int. Stat. Rev. *72*, 285–305.

https://doi.org/10.1111/j.1751-5823.2004.tb00237.x.
38. Grambsch, P.M., and Therneau, T.M. (1994). Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. Biometrika *81*, 515–526. https://doi.org/10.2307/2337123.
39. Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. Roy. Stat. Soc. B *57*, 289–300.
40. Benjamini, Y. (2010). Discovering the false discovery rate. J. Roy. Stat. Soc. B *72*, 405–416.
41. Quanjer, P.H., Stanojevic, S., Cole, T.J., Baur, X., Hall, G.L., Culver, B.H., Enright, P.L., Hankinson, J.L., Ip, M.S.M., Zheng, J., et al. (2012). Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. Eur. Respir. J. *40*, 1324–1343. https://doi.org/10.1183/09031936.00080312.
42. Hankinson, J.L., Odencrantz, J.R., and Fedan, K.B. (1999). Spirometric reference values from a sample of the general U.S. population. Am. J. Respir. Crit. Care Med. *159*, 179–187. https://doi.org/10.1164/ajrccm.159.1.9712108.
43. Liou, T., and Kanner, R. (2009). Spirometry. Clin. Rev. Allergy Immunol. *37*, 137–152. https://doi.org/10.1007/s12016-009-8128-z.
44. Cox, D.R. (1958). Some Problems Connected with Statistical Inference. Ann. Math. Stat. *29*, 357–372.
45. Majchrzak-Gorecka, M., Majewski, P., Grygier, B., Murzyn, K., and Cichy, J. (2016). Secretory leukocyte protease inhibitor (SLPI), a multifunctional protein in the host defense response. Cytokine Growth Factor Rev. *28*, 79–93. https://doi.org/10.1016/j.cytogfr.2015.12.001.
46. Chandler, J.D., Margaroli, C., Horati, H., Kilgore, M.B., Veltman, M., Liu, H.K., Taurone, A.J., Peng, L., Guglani, L., Uppal, K., et al. (2018). Myeloperoxidase oxidation of methionine associates with early cystic fibrosis lung disease. Eur. Respir. J. *52*, 1801118. https://doi.org/10.1183/13993003.01118-2018.
47. Rao, N.V., Marshall, B.C., Gray, B.H., and Hoidal, J.R. (1993). Interaction of secretory leukocyte protease inhibitor with proteinase-3. Am. J. Respir. Cell Mol. Biol. *8*, 612–616.
48. Vandooren, J., Goeminne, P., Boon, L., Ugarte-Berzal, E., Rybakin, V., Proost, P., Abu El-Asrar, A.M., and Opdenakker, G. (2018). Neutrophils and Activated Macrophages Control Mucosal Immunity by Proteolytic Cleavage of Antileukoproteinase. Front. Immunol. *9*, 1154. https://doi.org/10.3389/fimmu.2018.01154.
49. Liou, T.G., and Campbell, E.J. (1996). Quantum proteolysis resulting from release of single granules by human neutrophils: a novel, nonoxidative mechanism of extracellular proteolytic activity. J. Immunol. *157*, 2624–2631.
50. Liou, T.G., and Campbell, E.J. (1995). Nonisotropic enzyme–inhibitor interactions: a novel nonoxidative mechanism for quantum proteolysis by human neutrophils. Biochemistry *34*, 16171–16177.
51. Voynow, J.A., and Shinbashi, M. (2021). Neutrophil Elastase and Chronic Lung Disease. Biomolecules *11*, 1065. https://doi.org/10.3390/biom11081065.
52. Sagel, S.D., Thompson, V., Chmiel, J.F., Montgomery, G.S., Nasr, S.Z., Perkett, E., Saavedra, M.T., Slovis, B., Anthony, M.M., Emmett, P., and Heltshe, S.L. (2015). Effect of treatment of cystic fibrosis pulmonary exacerbations on systemic inflammation. Ann. Am. Thorac. Soc. *12*, 708–717. https://doi.org/10.1513/AnnalsATS.201410-493OC.
53. Sagel, S.D., Khan, U., Jain, R., Graff, G., Daines, C.L., Dunitz, J.M., Borowitz, D., Orenstein, D.M., Abdulhamid, I., Noe, J., et al. (2018). Effects of an Antioxidant-enriched Multivitamin in Cystic Fibrosis. A Randomized, Controlled, Multicenter Clinical Trial. Am. J. Respir. Crit. Care Med. *198*, 639–647. https://doi.org/10.1164/rccm.201801-0105OC.
54. Aratani, Y. (2018). Myeloperoxidase: Its role for host defense, inflammation, and neutrophil function. Arch. Biochem. Biophys. *640*, 47–52. https://doi.org/10.1016/j.abb.2018.01.004.
55. Hunt, W.R., Helfman, B.R., McCarty, N.A., and Hansen, J.M. (2016). Advanced glycation end products are elevated in cystic fibrosis-related diabetes and correlate with worse lung function. J. Cyst. Fibros. *15*, 681–688. https://doi.org/10.1016/j.jcf.2015.12.011.
56. Doll, R., and Hill, A.B. (1950). Smoking and carcinoma of the lung; preliminary report. Br. Med. J. *2*, 739–748.
57. Makam, M., Diaz, D., Laval, J., Gernez, Y., Conrad, C.K., Dunn, C.E., Davies, Z.A., Moss, R.B., Herzenberg, L.A., Herzenberg, L.A., and Tirouvanziam, R. (2009). Activation of critical, host-induced, metabolic and stress pathways marks neutrophil entry into cystic fibrosis lungs. Proc. Natl. Acad. Sci. USA *106*, 5779–5783. https://doi.org/10.1073/pnas.0813410106.
58. McMorran, B.J., Patat, S.A.O., Carlin, J.B., Grimwood, K., Jones, A., Armstrong, D.S., Galati, J.C., Cooper, P.J., Byrnes, C.A., Francis, P.W., et al. (2007). Novel neutrophil-derived proteins in bronchoalveolar lavage fluid indicate an exaggerated inflammatory response in pediatric cystic fibrosis patients. Clin. Chem. *53*, 1782–1791. https://doi.org/10.1373/clinchem.2007.087650.
59. Reynolds, P.R., Kasteler, S.D., Cosio, M.G., Sturrock, A., Huecksteadt, T., and Hoidal, J.R. (2008). RAGE: developmental expression and positive feedback regulation by Egr-1 during cigarette smoke exposure in pulmonary epithelial cells. Am. J. Physiol. Lung Cell Mol. Physiol. *294*, L1094–L1101. https://doi.org/10.1152/ajplung.00318.2007.
60. Ligthart, S., Sedaghat, S., Ikram, M.A., Hofman, A., Franco, O.H., and Dehghan, A. (2014). EN-RAGE: a novel inflammatory marker for incident coronary heart disease. Arterioscler. Thromb. Vasc. Biol. *34*, 2695–2699. https://doi.org/10.1161/ATVBAHA.114.304306.
61. Isoyama, N., Leurs, P., Qureshi, A.R., Bruchfeld, A., Anderstam, B., Heimburger, O., Bárány, P., Stenvinkel, P., and Lindholm, B. (2015). Plasma S100A12 and soluble receptor of advanced glycation end product levels and mortality in chronic kidney disease Stage 5 patients. Nephrol. Dial. Transplant. *30*, 84–91. https://doi.org/10.1093/ndt/gfu259.
62. Wittkowski, H., Sturrock, A., van Zoelen, M.A.D., Viemann, D., van der Poll, T., Hoidal, J.R., Roth, J., and Foell, D. (2007). Neutrophil-derived S100A12 in acute lung injury and respiratory distress syndrome. Crit. Care Med. *35*, 1369–1375. https://doi.org/10.1097/01.CCM.0000262386.32287.29.
63. Sagel, S.D., Kapsner, R.K., and Osberg, I. (2005). Induced sputum matrix metalloproteinase-9 correlates with lung function and airway inflammation in children with cystic fibrosis. Pediatr. Pulmonol. *39*, 224–232. https://doi.org/10.1002/ppul.20165.
64. Rubin, J.L., O'Callaghan, L., Pelligra, C., Konstan, M.W., Ward, A., Ishak, J.K., Chandler, C., and Liou, T.G. (2019). Modeling long-term health outcomes of patients with cystic fibrosis homozygous for F508del-CFTR treated with lumacaftor/ivacaftor. Ther. Adv. Respir. Dis. *13*, 1753466618820186. https://doi.org/10.1177/1753466618820186.
65. Bayfield, K.J., Douglas, T.A., Rosenow, T., Davies, J.C., Elborn, S.J., Mall, M., Paproki, A., Ratjen, F., Sly, P.D., Smyth, A.R., et al. (2021). Time to get serious about the detection and monitoring of early lung disease in cystic fibrosis. Thorax *76*, 1255–1265. https://doi.org/10.1136/thoraxjnl-2020-216085.
66. Reznikov, L.R. (2017). Cystic Fibrosis and the Nervous System. Chest *151*, 1147–1155. https://doi.org/10.1016/j.chest.2016.11.009.
67. Liou, T.G. (2019). The Clinical Biology of Cystic Fibrosis Transmembrane Regulator Protein: Its Role and Function in Extrapulmonary Disease. Chest *155*, 605–616. https://doi.org/10.1016/j.chest.2018.10.006.
68. Marshall, B.C., Butler, S.M., Stoddard, M., Moran, A.M., Liou, T.G., and Morgan, W.J. (2005). Epidemiology of cystic fibrosis-related diabetes. J. Pediatr. *146*, 681–687. https://doi.org/10.1016/j.jpeds.2004.12.039.
69. McGrath, L.T., Mallon, P., Dowey, L., Silke, B., McClean, E., McDonnell, M., Devine, A., Copeland, S., and Elborn, S. (1999). Oxidative stress during acute respiratory exacerbations in cystic fibrosis. Thorax *54*, 518–523. https://doi.org/10.1136/thx.54.6.518.
70. Conese, M., Castellani, S., Lepore, S., Palumbo, O., Manca, A., Santostasi, T., Polizzi, A.M., Copetti, M., Di Gioia, S., Casavola, V., et al. (2014). Evaluation of Genome-Wide Expression Profiles of Blood and Sputum Neutrophils in Cystic Fibrosis Patients Before and After Antibiotic Therapy. PLoS One *9*, e104080. https://doi.org/10.1371/journal.pone.0104080.
71. Laguna, T.A., Williams, C.B., Brandy, K.R., Welchlin-Bradford, C., Moen, C.E., Reilly, C.S., and Wendt, C.H. (2015). Sputum club cell protein concentration is associated with pulmonary exacerbation in cystic fibrosis. J. Cyst. Fibros. *14*, 334–340. https://doi.org/10.1016/j.jcf.2014.10.002.
72. Glasser, N.R., Hunter, R.C., Liou, T.G., and Newman, D.K.; Mountain West CF Consortium Investigators (2019). Refinement of metabolite detection in cystic fibrosis sputum reveals heme correlates with lung function decline. PLoS One *14*, e0226578. https://doi.org/10.1371/journal.pone.0226578.
73. Porro, C., Lepore, S., Trotta, T., Castellani, S., Ratclif, L., Battaglino, A., Di Gioia, S., Martínez, M.C., Conese, M., and Maffione, A.B. (2010). Isolation and characterization of microparticles in sputum from cystic fibrosis patients. Respir. Res. *11*, 94. https://doi.org/10.1186/1465-9921-11-94.
74. Forrest, O.A., Dobosh, B., Ingersoll, S.A., Rao, S., Rojas, A., Laval, J., Alvarez, J.A., Brown, M.R., Tangpricha, V., and Tirouvanziam, R. (2022). Neutrophil-derived extracellular vesicles promote feed-forward inflammasome signaling in cystic fibrosis airways. J. Leukoc. Biol. *112*, 707–716. https://doi.org/10.1002/JLB.3AB0321-149R.
75. Fujita, Y., Kosaka, N., Araya, J., Kuwano, K., and Ochiya, T. (2015). Extracellular vesicles in

lung microenvironment and pathogenesis. Trends Mol. Med. *21*, 533–542. https://doi.org/10.1016/j.molmed.2015.07.004.

76. Genschmer, K.R., Russell, D.W., Lal, C., Szul, T., Bratcher, P.E., Noerager, B.D., Abdul Roda, M., Xu, X., Rezonzew, G., Viera, L., et al. (2019). Activated PMN Exosomes: Pathogenic Entities Causing Matrix Destruction and Disease in the Lung. Cell *176*, 113–126.e15. https://doi.org/10.1016/j.cell.2018.12.002.
77. Shmueli, G. (2010). To Explain or to Predict? Stat. Sci. *25*, 289–310. https://doi.org/10.1214/10-STS330.
78. Cox, D.R. (1958). Planning of Experiments (Wiley).
79. Nichols, D.P., Paynter, A.C., Heltshe, S.L., Donaldson, S.H., Frederick, C.A., Freedman, S.D., Gelfond, D., Hoffman, L.R., Kelly, A., Narkewicz, M.R., et al. (2022). Clinical Effectiveness of Elexacaftor/Tezacaftor/Ivacaftor in People with Cystic Fibrosis: A Clinical Trial. Am. J. Respir. Crit. Care Med. *205*, 529–539. https://doi.org/10.1164/rccm.202108-1986OC.
80. Harris, P.A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., and Conde, J.G. (2009). Research electronic data capture (REDCap)–a metadata-driven methodology and workflow process for providing translational research informatics support. J. Biomed. Inf. *42*, 377–381. https://doi.org/10.1016/j.jbi.2008.08.010.
81. Fuchs, H.J., Borowitz, D.S., Christiansen, D.H., Morris, E.M., Nash, M.L., Ramsey, B.W., Rosenstein, B.J., Smith, A.L., and Wohl, M.E.; The Pulmozyme Study Group (1994). Effect of Aerosolized Recombinant Human DNase on Exacerbations of Respiratory Symptoms and on Pulmonary Function in Patients with Cystic Fibrosis. N. Engl. J. Med. *331*, 637–642. https://doi.org/10.1056/NEJM199409083311003.
82. Davison, A.C., and Hinkley, D.V. (1997). Bootstrap Methods and Their Application, 1st ed. (Cambridge University Press).
83. Standardization of Spirometry (1994). Update. American Thoracic Society (1995). Am. J. Respir. Crit. Care Med. *152*, 1107–1136. https://doi.org/10.1164/ajrccm.152.3.7663792.
84. Miller, M.R., Hankinson, J., Brusasco, V., Burgos, F., Casaburi, R., Coates, A., Crapo, R., Enright, P., van der Grinten, C.P.M., Gustafsson, P., et al. (2005). Standardisation of spirometry. Eur. Respir. J. *26*, 319–338. https://doi.org/10.1183/09031936.05.00034805.
85. Tukey, J.W. (1977). Exploratory Data Analysis (Addison-Wesley Publishing Company).
86. R Core Team (2020). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).
87. Venables, W.N., and Ripley, B.D. (2010). Modern Applied Statistics with S (Springer).
88. Dasenbrook, E.C., Checkley, W., Merlo, C.A., Konstan, M.W., Lechtzin, N., and Boyle, M.P. (2010). Association between respiratory tract methicillin-resistant Staphylococcus aureus and survival in cystic fibrosis. JAMA *303*, 2386–2392. https://doi.org/10.1001/jama.2010.791.
89. van Buuren, S., Boshuizen, H.C., and Knook, D.L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. Stat. Med. *18*, 681–694. https://doi.org/10.1002/(SICI)1097-0258(19990330)18:6<681::AID-SIM71>3.0.CO;2-R.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-human CRP (2.5 μg/ml)	Abcam, Cambridge, MA, USA www.abcam.com	Abcam Cat# ab31156; RRID AB_2085618
Mouse monoclonal anti-human Calprotectin (1.5 μg/ml)	LS-Bio, Shirley, MA, USA www.lsbio.com	LS-Bio (LifeSpan) Cat# LS-C96223; RRID AB_10567208
Rabbit polyclonal anti-human HMGB1 (1.5 μg/ml)	Upstate Biotechnology (Now Millipore, Burlington, MA, USA www.emdmillipore.com/US)	Upstate Cat# 07-584; RRID AB_11210378
Mouse monoclonal IgG anti-human CRP (1.5 μg/ml)	Abcam, Waltham, MA, USA www.abcam.com	Abcam Cat# ab136176; RRID AB_2747847
Rabbit polyclonal anti-human Calprotectin (1:1500)	LS-Bio, Shirley, MA, USA www.lsbio.com	LSBio (LifeSpan) Cat# LS-C122793-20, RRID:AB_10805991
Mouse monoclonal anti-human HMGB1 (0.75 μg/ml FC)	R&D Systems, Minneapolis, MN, USA www.rndsystems.com	R&D Systems Cat# MAB1690; RRID AB_2117897
Goat Anti-Mouse Polyclonal IgG-FC 1:2000	Millipore, Burlington, MA, USA www.emdmillipore.com/US	Millipore Cat# AP127P; RRID AB_92472
Goat Anti-Rabbit IgG-FC 1:2000	Santa Cruz Biotechnology, Dallas, TX, USA www.scbt.com	Santa Cruz Biotechnology Cat# SC-2004; RRID AB_631746
Goat Anti-Mouse Polyclonal IgG (FC 1:2000)	Millipore, Burlington, MA, USA www.emdmillipore.com/US	Millipore Cat# AP127P; RRID AB_92472
Chemicals, peptides, and recombinant proteins		
1% bovine serum albumin	MilliporeSigma, Burlington, MA, USA www.emdmillipore.com/US	Cat# 9048-46-8
3,3′,5,5′-tetramethylbenzidine substrate solution	Thermo Scientific, Waltham, MA, USA www.thermofisher.com	Cat# N301
10% newborn calf serum in phosphate buffered saline	MilliporeSigma, Burlington, MA, USA www.emdmillipore.com	Cat# 16010142
C-Reactive Protein (CRP)	Abcam, Waltham, MA, USA www.abcam.com	Cat# ab167710
Calprotectin or S100A8/A9 Heterodimer	Biolegend, San Diego, CA, USA www.biolegend.com	Cat# 753404
Hanks Buffered Saline Solution	Sigma-Aldrich, Burlington, MA, USA www.sigmaaldrich.com	Cat# H9269
High Mobility Group Box-1 (HMGB1)	Sigma-Aldrich, Burlington, MA, USA www.emdmillipore.com	Cat# H4652
Roche cOmplete™ Protease Inhibitor Cocktail tablets	Sigma-Aldrich, Burlington, MA, USA www.emdmillipore.com	Cat# 11697498001
Streck Cell Preservative Solution	Streck, La Vista, Nebraska, USA	Cat# 213355
Critical commercial assays		
Neutrophil elastase activity	The Pediatric Clinical Translational Research Center Core Laboratory, Children's Hospital Colorado	N/A
Custom Luminex 4-Plex Kit	Biotechne, Minneapolis, MN, USA www.rndsystems.com	Cat# LXSAHM-4
Custom Luminex 16-Plex Kit	Biotechne, Minneapolis, MN, USA www.rndsystems.com	Cat# LXSAHM-16

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Theodore G. Liou, MD (ted.liou@utah.edu).

Materials availability

New materials generated by this report include remaining lipid, aqueous and pellet fractions of the original sputum samples. Please contact Dr. Liou to inquire about sharing fully annotated deidentified samples. Any transfer requires execution of a Materials Transfer Agreement with the University of Utah and may require a fee for processing, handling and shipping.

Data and code availability

- The data reported in this paper will be shared by the lead contact in deidentified form upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Ethical reviews

Following finalization of study protocols, we submitted the study to the ethical review boards or committees at each of the nine participating MWCFC centers (Table S1). No study procedures were started prior to approval at each of the centers.

Study participants

This was a prospective observational study examining airway inflammation and any relationships with time to the next pulmonary exacerbation of CF in participants who were clinically stable at the time of enrollment and able to produce expectorated sputum. We recruited participants at random from the nine Mountain West CF Centers accredited by the US CF Foundation (Bethesda, Maryland). We followed participants from enrollment for a minimum of 18 months up to 2.65 years when the study concluded.

Participants were required to be 12 years of age or older, and we selected patients after block randomization for two groups, age 12 to <18 and 18 or older (see below for randomization details). We excluded children with CF younger than 12 because of their inconsistent ability to expectorate sputum. We made no specification on race, ethnicity or sex, but the population recruited reflects characteristics of the population of people with CF in the CF Foundation Patient Registry for 2014.[33] Table 1 presents the characteristics of the participants. The proportion of participants who were female was 0.46. Sex was used as an adjustment in all analyses (see below) but was not reported in our main results because it had no statistical impact. Race and ethnicity numbers are withheld as all groups that were not white each numbered less than 5. Race and ethnicity were used in estimations of FEV_1% along with age, sex and height. Socioeconomic status was not recorded.

METHOD DETAILS

Pre-study preparations and personnel training

Development of all study procedures was initiated at the University of Utah and reviewed, modified and finally approved by all participating MWCFC centers. Patients at the University of Utah provided feedback and suggestions on experimental procedures directly involving study participants. University of Utah personnel tested all sputum collection and processing procedures as well as shipping procedures on actual samples collected with informed consent from 10 participants prior to the main study. (These samples were not included in the current work.) Evaluations of laboratory results of shipping led to a change in procedure from central processing of all samples after overnight wet-ice shipping to on site sample processing with later batch shipping on dry ice, the method underlying results reported here.[33]

All participating MWCFC personnel were trained at the University of Utah on study background, goals, inclusion and exclusion criteria, good clinical practice, patient safety and study procedures, data entry using the Research Electronic Data Capture (REDCap) system[80] and records security. The initial University of Utah team provided teaching on the finalized clinic and laboratory protocols at the meeting, at subsequent onsite study initiation meetings and at additional in person trainings at the University of Utah as requested by MWCFC site personnel.

We produced a three part video introduction to the study design underlying the project as an interview of Sir David Cox by Ruth Keogh and a training video on sputum collection and processing with the University of Utah Research Team that remain available for viewing at https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-019-0705-0. Specifically, see Additional Files 1-3 for the interview, Additional File 4 for a transcript and Additional File 7 for sputum processing.[33]

Once finalized, the study background, purpose, procedures, funding sources, informed consent and assent documents, training documentation for work with human subjects were submitted to the investigational review boards or equivalent ethical review committees at each of the MWCFC Centers (Table S1). No study activities involving participants were begun until formal approval at each of the centers was received. Changes to any aspect of the study were submitted to each approving entity by amendment and were not implemented until approved.

Study design, participants and sample size

To mitigate bias, we randomly selected participants from MWCFC care center (Table S1) patient lists older than 12 years able to expectorate sputum for prospective observational study of biomarkers of airway inflammation.[33] After successful institutional reviews (Table S1), each participating center identified candidates for study who were 18 years or older, able to provide informed consent and able to expectorate sputum. Each center identified candidates who were older than 12 up to 18 years of age who were able to provide assent, had parents or guardians who could provide informed consent and were able to expectorate sputum or tolerate sputum induction as part of routine care. We generated two blocks of randomly chosen letters of the alphabet for adolescent and adult candidates to avoid over representation of either group.

For each recruiting center, we assigned a threshold letter for enrollment. Threshold letters for centers were chosen to allow enrollment from all centers proportional to their patient population sizes within the MWCFC and allow for a 10% combined clinic no-show and study refusal rate. We adjusted the center threshold letters after the first month and after each quarter of the year during the study to adjust enrollment rates to prevent over representation of patients from any one quarter of the year and maintain proportional enrollments among the participating centers as much as feasible. These adjustments did not affect the randomized nature of enrollment but may have delayed enrollment for a few participants. The pre-clinic visit assignments eliminated the need for randomization for enrollment by a research coordinator at a clinical appointment time but excluded any new patients to the clinics from participation once the study was begun.

Candidates who came to clinic who were clinically stable and had a personal randomly assigned letter earlier in the alphabet than the assigned center letter were approached for enrollment. After we obtained written informed consent from adults and assent from adolescents with informed parental or guardian consent, we enrolled clinically stable participants from December 8, 2014 through January 16, 2016. All consents were obtained in person.

Our primary outcome was days from enrollment to next pulmonary exacerbation requiring hospitalization for acute treatment of at least one symptom and one sign of CF pulmonary exacerbation.[17] Symptoms included increased sputum, cough or dyspnea, chest pain or tightness, participant reported hemoptysis, fever, chills, arthralgias, fatigue. Signs included 10% drop in FEV_1 or forced vital capacity, temperature > 38.4°C, healthcare personnel witnessed hemoptysis > 100 mL per episode, SaO_2 < 90% or PaO_2 < 60 mm Hg despite usual oxygen, a drop in SaO_2 of 5% in adolescents, increased supplemental oxygen and unplanned weight loss $\geq$ 5% of baseline body weight over 3 months. We chose this definition of an exacerbation because it may be applied prospectively and requires no retrospective adjudication based on use of antibiotics or other standard responses indicating an exacerbation as in many prior studies.[81]

We allowed site principal investigator (**PI**) discretion for symptoms and signs not listed here, which was exercised once during the study. We allowed for a site PI to consider admission for respiratory arrest as an exacerbation of CF, but no participants were so unfortunate. The days to the next pulmonary exacerbation were not affected by the success of sputum collection at the next exacerbation. In less than 5 cases, participants were admitted for their next exacerbation at non-MWCFC institutions. Research coordinators obtained and reviewed charts in those cases to ascertain the occurrence and timing of the next exacerbation and verified findings with the appropriate MWCFC Site PI.

We selected a total number of participants to enroll based on power calculations derived from prior experience with biomarker studies with people with CF and pulmonary exacerbations.[33] Using data including age, number of exacerbations in the year prior to enrollment, FEV_1%, HMGB1 sputum concentrations and time to next exacerbation from our prior work for 26 participants who were clinically stable,[17] we sampled 20, 26, 40, 60, 80, 100, 125 and 150 participants with replacement to create synthetic data sets from among the 26 actual patients. We bootstrapped[82] Cox proportional hazards models[36] with 1000 iterations per synthetic data set. We modeled time to next exacerbation as the outcome and used HMGB1 concentration as the main independent variable. We repeated the procedure for models adjusted by age, sex, number of prior pulmonary exacerbations within 1 year and FEV_1% that was itself calculated from FEV_1, age, sex, height, race and ethnicity using NHANES III equations.[42] We used the percentage of models within each 1000 bootstrapped model set with $p \leq 0.01$ or 0.05 to derive the power of a data set to reach an α of 0.01 or 0.05 for the size of each synthetic data set. For HMGB1, the results for 40 participants revealed 90% power for $\alpha = 0.01$ (meaning that 90% of the 1000 Cox proportional hazards models run with datasets of 40 simulated patients had $p \leq 0.01$).

We performed similar bootstrapped power calculations for other biomarkers that we previously studied.[17] For GMCSF sputum concentration as the independent variable and the acute decrease in FEV_1% accompanying a diagnosis of pulmonary exacerbation as the dependent variable, we used linear regression with up to 175 participants in synthetic data sets. We found that 125 participants implied 85% power for $\alpha = 0.05$ while 175 participants implied 80% power for $\alpha = 0.01$. Based on these calculations, we targeted an enrollment of 125 to 175 participants as the maximum number of participants to enroll.

Sputum and biomarker collections and measurements

We processed, aliquoted and froze expectorated sputum immediately after collection (mean = 54, SD = 98, max = 225 minutes on ice) to allow batch laboratory evaluations to reduce variation in results due to artifacts related to assays performed at different times.[33] We collected 114 enrollment samples when participants were clinically stable using sterile 50 mL conical bottom centrifuge tubes. We collected 52 samples within 48 hours of the diagnosis of the next pulmonary exacerbation; 32 convalescence samples within 4 to 12 weeks after exacerbation diagnosis; 12 samples within 48 hours of any subsequent exacerbation; and 55 samples at study end. We annotated each sample with clinical information, cell counts and differentials.[33]

For this work, we used only the 114 enrollment samples because we sought to understand relationships between inflammatory biomarkers observed during clinical stability and the next pulmonary exacerbation from the viewpoint of a clinician seeing a patient in clinic for routine

follow up. Such a clinician would not have access to information from future collections of data or sputum measurements. Because of the focus on biomarkers previously observed to have relationships of some kind with past, concurrent or future exacerbations that were measured from aqueous sputum fractions, we used only the aqueous fractions obtained after centrifugation of well mixed samples that were carefully separated from top lipid and bottom pellet fractions.[33]

We collected samples on ice and transported from each MWCFC clinic to its respective laboratory for processing. All samples were processed within 4 hours, but we specified that sputum expectoration time was limited to 20 minutes, transport on ice was limited to 40 minutes and post-collection processing was to begin no later than 60 minutes after the start of collection. Site visits by the University of Utah team verified adherence to these guidelines. Each sample was labeled at the point of collection to identify the participant and the type of sample relative to study stage or exacerbation.

In the lab, samples were diluted 1:1 with Hanks Buffered Saline Solution (key resources table) and vortex mixed for 1 minute. Using a sterile disposable pipette, 0.25 mL of the mixed sample was transferred to a 1.8 mL tube containing 0.25 mL of Streck Solution (key resources table). The Streck mixed sample was sent via Fedex (Memphis, TN) to the University of Utah for cell counts and differentials. Cell counts were performed manually with a hemocytometer and desktop light microscope. Differentials were determined after modified Wright Staining of smears. Cells were not subjected to cytospinning due to bias associated with increased fragility of some leukocytes which cannot be counted if they rupture, however, this procedure makes differentiating between cell types more difficult.

The remaining HBSS diluted sample was centrifuged at 2,800 g at 4°C for 20 minutes. Top lipid, middle aqueous and bottom pellet layers were seperated using sterile transfer pipettes to avoid contamination of the aqueous layer by the other two layers. The lipid layer was frozen in a labeled container and remains unused. The aqueous layer was divided in two with the first half diluted 1:1 with additional HBSS, vortex mixed for 10 s and frozen as multiple 1 mL aliqouts. The second half was diluted 1:1 with protease inhibitor cocktail (key resources table), vortex mixed 10 s and frozen as 1 mL aliquots. Pellet fractions were frozen without further processing. All fractions were frozen at -70°C until assayed. At least 2 aliquots of each fraction were frozen, and samples were shipped on dry ice in two batches per patient to avoid losses due to shipping failures. We encouraged samples from more than 1 patient to be shipped together. This procedure prevented loss of any single sample despite thawed shipments on two occasions.[33]

Concurrent with sample collections, we measured forced expiratory volume in 1 s ($\mathbf{FEV_1}$) in accordance with American Thoracic Society and European Respiratory Society guidelines,[83,84] and we estimated percent predicted FEV_1 (FEV_1%) based on age, height and sex at the time of testing.[42,84] A few individuals enrolled in the study belong to groups for which no equations are published (number of individuals and nature of groups withheld to protect privacy). We used a selection of equations to generate different estimates of FEV_1% to enable inclusion of these individuals in all analyses.[43] We based prior year exacerbation counts on hospitalization dates within one year before enrollments. We calculated weight-for-age *z*-score and 5-year prognostic scores (higher predicts longer survival).[29] Prognostic scores were recalculated each time we used a different set of FEV_1% estimating equations.

The Pediatric Clinical Translational Research Center Core Laboratory at Children's Hospital Colorado and University of Colorado Anschutz Medical Campus (Aurora, CO, USA) spectrofluorometrically measured NE activity[18] using aqueous fractions frozen without protease inhibitor cocktail. Prior literature almost exclusively uses NE activity as the target of study while for other enzymes in our study (MPO, PR3, MMP9), enzyme linked immunosorbent assay (ELISA) results are used for analysis. The University of Utah CF Center Laboratory (Salt Lake City, UT, USA) measured calprotectin, CRP, and HMGB1 by ELISA (key resources table) using aqueous fraction aliquots frozen with protease inhibitor cocktail. We shipped frozen aqueous samples with protease inhibitor cocktail (Fedex, Memphis, TN, USA) to R&D Systems' Biomarker Testing Service (Biotechne, Minneapolis, MN, USA) for remaining assays. Because of different standard curves and dilution requirements, Biotechne assayed MMP9, MPO, PR3 and ENRAGE using a 4-plex human luminex kit (R&D Systems Cat # LXSAHM-04) after 1:50 dilution. The remaining 16 biomarkers (GMCSF, ICAM1, IFNγ, IL1β, IL5, IL6, IL8, IL10, IL17A, S100A8, S100A9, SLPI, sRAGE, TARC, TNFα, YKL40) were assayed using a 16-plex human luminex kit (R&D Systems Cat # LXSAHM-16) after 1:2 dilution.

Because multiplex assays require a single dilution of each sample, optimization across all biomarkers resulted in some single biomarker values outside the specific standard curve range for limits of detection. After considering standard curve values for each biomarker, we used quantile plots to visually ascertain[85] upper and lower limits of detection for each biomarker. In order to retain those values which provided partial information for use in the study, we adjusted values outside the limits of detection. Concentrations outside the limits of detection were assigned values 0.1% above upper or below lower detection limits to retain partial information prior to natural logarithmic transformation and use in statistical analysis. The procedure avoids creating outlier values that can influence subsequent analyses such as arbitrarily assigning a log-transformed value of 1 when values within the limits of detection have, for example, a mean of 14.4 (SD = 1.32) for MMP9 (Table S6).

ELISA procedures for measurements of calprotectin, CRP and HMGB1

We incubated 96-well plates (Costar, Corning Inc, Costar, NY, USA) overnight with capture antibody diluted according to manufacturer recommendations or 1% bovine serum albumin (BSA, MilliporeSigma, Burlington, MA, USA) or 10% newborn calf serum in phosphate buffered saline (NCS-PBS, MilliporeSigma). Plates were washed 4 times with PBS prior to addition of 1st antibodies then washed 6 times prior to addition of horse radish peroxidase-conjugated 2nd antibodies. Standards were diluted 1:2 for standard curves using manufacturer's recommendations or 1% BSA or 10% NCS-PBS (MilliporeSigma). All assays were incubated with 3,3′,5,5′-tetramethylbenzidine (TMB) substrate solution (Thermo Scientific, Waltham, MA, USA). Reactions were stopped using 0.18 M H_2SO_4 after 15-30 minutes and read in an ELISA plate reader at OD_{450}. Standard curves were constructed using linear regression of log-transformed mean fluorescence intensities and log-transformed known protein concentrations.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical software

We use the R statistical programming environment[86] for data handling and analysis. We use the basic packages included in the stable release (version 4.3.x) along with the survival package which includes generalized linear model (for binomial, gaussian and quasi-poisson regression) and Cox proportional hazards model functions, the future and future_lapply packages for parallel processing and the rlist package for convenient handling of list outputs. We generally set $p < 0.05$ as the threshold for statistical significance, but we performed false discovery rate[39] analysis using thresholds of 0.001, 0.01, 0.05, 0.1 and 0.2 as illustrated in figures. All figures were drawn using standard plotting commands provided in base R software.

Statistical analysis procedures

We performed causal inference using key principles suggested for control of confounding.[30] We performed purposeful variable selection and postulated a specific causal model for biomarkers, clinical variables, and outcome using a directed acyclic graph (Figure 1). We evaluated Pearson product moment correlations among detection limit adjusted biomarker measurements and used family-wise error correction with $\alpha \leq 1 \times 10^{-5}$ to reduce the number of reported correlations to an easily comprehensible number.[87] We estimated the distribution of time to next pulmonary exacerbation by Kaplan-Meier[34] stratified by the nine clinical survival predictors[14,29] (see introduction), followed by logistic regression[35] and proportional hazards model[36] evaluations using clinical variables as explanatory. We examined proportional hazards models for departures from the assumption of proportionality by examining the χ-square distribution of weighted Schoenfeld residuals.[38] We used exacerbation within the first follow up year as the logistic regression outcome and time to next exacerbation as the proportional hazards outcome.

Intrinsic variables such as age, sex, and prior pulmonary exacerbations cannot mediate but may confound inflammatory effects for future events thus may adjust time to next pulmonary exacerbation models.[30] In contrast, variables that describe active disease such as nutritional status, diabetes and lung function potentially mediate future pulmonary exacerbations.

Assessments of confounding and mediation

We assessed confounding and mediation (Figure 1) in steps (Figure S1).[31] First, we fitted a regression model for each biomarker as a univariable with 5-year prognostic risk score variables[14,29] (age, sex, FEV_1%, weight-for-age *z*-score and status of diabetes, pancreatic insufficiency, MSSA and BCC infections) and other potential confounders or mediators as the outcomes (thin arrows on left side of Figure 1). We used linear regression for age, vital signs, FEV_1%, weight-for-age *z*-score and total cell count; quasi-Poisson regression for prior year pulmonary exacerbation count; and logistic regression for status of pancreatic sufficiency, diabetes and MSSA and BCC infections. Because of clinical importance,[13,88] we additionally evaluated methicillin resistant *Staphylococcus Aureus* (**MRSA**) and *Pseudomonas aeruginosa* status as outcome variables using logistic regression.

Second, we evaluated proportional hazards models for time to next pulmonary exacerbation with the potentially explanatory inflammatory biomarkers as independent univariables (thin arrows on right side of Figure 1). To each model, we added, one at a time, the variables that had relationships with the inflammatory biomarkers in the first part of the confounding and mediation analysis as independent adjustment variables. We identified variables that substantially reduced biomarker effect sizes and significance as mediators for exclusion from further models and considered remaining variables as potential confounders.

Finally, we considered a variable as potentially confounding and useful as an adjustment variable if it fulfilled two conditions. (1) It had significant associations with multiple biomarkers in regression models (linear, quasi-Poisson or logistic from the first step above). (2) It improved the fit of proportional hazards models of time to next exacerbation with each biomarker when it was included as an adjustment variable, especially for the same biomarkers for which it fulfilled condition (1).

Development of explanatory models and false discovery rate analyses

Following confounding and mediation assessments[31] of prognostic risk score variables[14,29] and selected clinically important variables such as MRSA, sputum neutrophil and total cell counts,[13,88] we developed explanatory models for time to next pulmonary exacerbation (thick arrow, Figure 1). We fitted proportional hazards models[36] for each biomarker, adjusted for confounding with exclusion of mediating variables.[31]

Because we evaluated 24 biomarkers and found 10 which had *p* values < 0.05 in proportional hazards models of time to next exacerbation with each biomarker as the independent variable adjusted by age and prior pulmonary exacerbations, we performed false discovery rate (**FDR**) analysis to reduce the chance of categorizing spurious associations as significant.[39] We chose a graphical approach because it is easy to evaluate multiple biomarker results simultaneously and illustrate the findings.

We selected 20% as the threshold rate of false discovery above which we would discourage further investigation of a biomarker. To perform the evaluation for our main results (see Figure 2, for example), we ranked biomarkers based on *p* values from the 24 proportional hazards models of time to next exacerbation, each adjusted by age and prior pulmonary exacerbations in the year prior to enrollment. We plotted the *p* values on the y-axis and *p*-value rank on the x-axis for the graphical analysis. We drew a line through the origin with a slope equal to FDR threshold/24 or 0.2/24. All biomarkers below and to the right of this line have an FDR < 20% or true discovery rate of $(1 - FDR) > 80\%$. We drew a second line for FDR threshold = 10%.

As illustrated in Figure 2, ENRAGE has a true discovery rate > 90%. YKL40 also appears to have a true discovery rate > 90%, however, the biomarkers plotted between ENRAGE and YKL40 have intervening ranks and fail the FDR threshold = 10% line. The first biomarker that crosses above and to the left of a threshold line is considered to disqualify any subsequent biomarkers, thus the biomarkers ranked 2-10 (Figure 2B) are considered all to have FDR < 20% or true discovery rate (*1 – FDR*) of > 80%. We did similar graphical FDR analyses for other analyses with some differences in chosen FDR thresholds (Figures 3 and S2–S7).

Sensitivity analyses

We assessed Cox proportional hazards model sensitivities to anti-inflammatory treatments including steroids, chronic antibiotics (oral azithromycin, inhaled aztreonam and tobramycin) and potentially anti-inflammatory CF transmembrane regulator protein (**CFTR**) modulators, ivacaftor and ivacaftor-lumacaftor.[13] We included these treatments as binary adjustment variables one at a time in the explanatory models described above for each biomarker and examined the results seeking substantial changes in the effect sizes for each biomarker and the adjustments for age and prior pulmonary exacerbation counts. Effect sizes that reversed sign (or hazard ratios that transitioned from less than 1 to more than 1 or vice versa) would have indicated sensitivity to a treatment as would a change in effect resulting in a clinically important change in time to next exacerbation. The latter threshold for sensitivity is necessarily harder to define, but a change in relationship of a biomarker with time to exacerbation measured in weeks or months which would be meaningful to a person with CF, for example, would have been considered evidence of model sensitivity.

We assessed the impact of substituting Global Lung Initiative (GLI) equations[41] to estimate $FEV_1\%$ for NHANES III equations.[42] We substituted $FEV_1\%_{GLI}$ for $FEV_1\%_{NHANESIII}$ and compared estimates, 95% CI and *p*-values for linear regressions of biomarkers with $FEV_1\%$ and hazard ratios, 95% CI and *p*-values for proportional hazards models of time to next exacerbation with $FEV_1\%$ and biomarkers. We would have considered a change in sign for estimates or a shift from below to greater than 1 (or vice versa) for hazard ratios or change in interpretation of $FEV_1\%$ as a mediator of biomarker effects as evidence of model sensitivity.

Low sample volumes prevented some measurements. We examined whether participants on ivacaftor or ivacaftor-lumacaftor produced less sputum by weight for study than participants not treated with a CFTR modulator and compared using a *t*-test. We performed three analyses seeking evidence that missingness was other than completely at random: using a dataset (1) deleting records with completely or partially unmeasured biomarkers, (2) deleting only records with completely unmeasured biomarkers and (3) retaining all records after multiple imputation by chained equations[89] for completely missing biomarker values.

ADDITIONAL RESOURCES

This is an observational study of airway inflammation without registration. The times of study planning (2012) and study initiation with first enrollment (2014) predate the allowance for observational trial registration at clinicaltrials.gov. Registration was not required because the study did not involve assignment of participants to any intervention nor did it primarily observe the outcomes of interventions given in the course of regular care.

The study design and sputum collection and processing protocols were previously published: Liou, T.G., Adler, F.R., Argel, N., Asfour, F., Brown, P.S., Chatfield, B.A., Daines, C.L., Durham, D., Francis, J.A., Glover, B. et al. (2019). Prospective multicenter randomized patient recruitment and sample collection to enable future measurements of sputum biomarkers of inflammation in an observational study of cystic fibrosis. BMC Med Res Methodol *19*, 88. https://doi.org/10.1186/s12874-019-0705-0. Teaching videos (with a transcript) to improve understanding of our study design incorporating randomized selection of participants and for laboratory processing of samples are included in this reference.

Website with step by step guidance on mediation analysis with examples:

Kenny, D.A. Mediation. https://davidakenny.net/cm/mediate.htm.