SUPPLEMENTARY MATERIAL

Risk factors for lung disease progression in children with cystic fibrosis

Statistical Analysis

Baseline characteristics were expressed as mean (standard deviation [SD]) or median (interquartile range [IQR]) for numerical variables, and as number (percentage) for categorical variables. For the longitudinal analysis of the first outcome measure, FEV_1%pred, linear mixed-effects models were used, in which the numerical variable time (0–5 years), potential risk factors (see above), and their interactions with time were entered as fixed factors. Since time is entered here as a numerical variable and measured in years, the effect of a potential risk factor on FEV1%pred indicates whether the annual change in FEV1%pred is affected by this factor. The interactions between factors and time were included to estimate the effect of the potential risk factors on the mean annual change in FEV_1%pred. A random intercept and random slope (time) with an unstructured covariance structure between intercept and slope were included in the model to deal with the dependency between repeated measures from the same patient.

For the longitudinal effect on future PEx (measured at years one, two, three, four, and five), generalised estimating equations (GEE) were used, where time and potential risk factors (including PEx at baseline) were added as fixed factors. Since no interactions between time and the potential risk factors were included, the effect of a potential risk factor is expressed as an overall odds ratio (OR), pooled over all time-points.. A compound symmetry working correlation matrix, using the model-based estimator was used for the random part, in which the unstructured covariance structure and/or robust estimator were considered as sensitivity analyses.

There were no restrictions on the minimum follow-up data required for the analysis. Multiple imputation was used to deal with missing data by creating 50 complete datasets. The maximum number of iterations was set equal to 20, where convergence was checked by inspecting the trace lines. The missing covariate values were imputed using all other baseline variables as predictors. Missing outcome values at a certain time-point were imputed using all baseline variables in addition to the patients' outcome measurements that were obtained at the other time-points. As for sensitivity analysis, the complete case analysis, in which the patients with one or more missing covariate values at baseline are excluded from the analysis, was also performed.

For both outcome measures, first univariable (explorative) models and thereafter multivariable models were assessed. The estimated effect (pooled for five years) of each potential risk factor on mean annual change in FEV_1%pred or overall odds ratio (OR) for PEx (over all time-points) was reported together with the corresponding 95% confidence interval (95% CI) and p-value. Multicollinearity of the models was checked (all variance inflation factors were <1.7). Two-sided p-values smaller than or equal to 0.05 were considered statistically significant. Baseline characteristics and GEE analyses were obtained using IBM SPSS Statistics for Windows (version 23.0). The multiple imputation part was performed using the MICE package in R (version 3.2.3). [1]

Table S1. Baseline characteristics PPI use versus no PPI use in baseline year

Characteristic	PPI use in baseline year (N=194)	No PPI use in baseline year (N=315)	p-value
Age, mean (SD)	11.7 (3.5) (N=194)	11.9 (3.9) (N=315)	0.518
Male sex, N (%)	94 (49) (N=192)	163 (52) (N=312)	0.474
F508del homozygous, N (%)	134 (69) (N=194)	167 (53) (N=315)	<0.001*
F508del heterozygous, N (%)	52 (27) (N=194)	122 (39) (N=315)	0.006*
BMI *z*-score, mean (SD)	-0.26 (0.88) (N=190)	-0.08 (0.92) (N=304)	0.033*
Best FEV_1%pred, mean (SD)	88.0 (18.6) (N=189)	88.9 (16.3) (N=304)	0.605
Best FVC%pred, mean (SD)	94.1 (15.7) (N=189)	88.9 (16.3) (N=304)	0.312
Use of pancreatic enzymes, N (%)	187 (97) (N=193)	252 (80) (N=314)	<0.001*
Use of prophylactic inhaled antibiotics, N (%)	74 (38) (N=194)	98 (31) (N=314)	0.109
Use of macrolide antibiotics, N (%)	68 (35) (N=194)	85 (27) (N=315)	0.054
Use of ICS, N (%)	58 (30) (N=194)	60 (19) (N=315)	0.005*
Colonisation with *Pseudomonas aeruginosa*, N (%)	53 (28) (N=187)	59 (20) (N=300)	0.027*
ABPA, N (%)	18 (9) (N=194)	33 (11) (N=315)	0.662
CFRD, N (%)	22 (11) (N=194)	15 (5) (N=314)	0.006*
Course(s) of IV antibiotics at baseline PEx, N (%)	62 (50) (N=125)	67 (36) (N=188)	0.014*

Numbers in brackets (N) refer to non-missing data. ABPA, allergic bronchopulmonary aspergillosis; BMI, body mass index; CFRD, cystic fibrosis related diabetes, FEV_1%pred, forced expiratory volume in 1 second as a percentage of predicted value; FVC%pred, forced vital capacity as a percentage of predicted value; ICS, inhaled corticosteroids; IV, intravenous; PEx, pulmonary exacerbations; PPI, proton pump inhibitors. *p<0.05

Table S2. Estimated adjusted effect on mean annual change of FEV_1% predicted, multivariable model for total cohort (N=545)

Risk factor	**Effect (95% CI)**	**p-value**
Male sex	0.07 (−0.48, 0.63)	0.793
Age	0.004 (−0.107, 0.116)	0.939
F508 del homozygous	−0.31 (−0.87, 0.25)	0.280
BMI *z*-score	0.018 (−0.335, 0.371)	0.921
FEV_1%pred	−0.005 (−0.044, 0.034)	0.849
Pancreatic enzymes	−0.09 (−1.12, 0.93)	0.858
Prophylactic inhaled antibiotics	−0.53 (−1.25, 0.19)	0.150
Prophylactic macrolide	0.09 (−0.64, 0.83)	0.803
PPI use	−0.69 (−1.26, −0.12)	0.017*
Inhaled corticosteroids	0.51 (−0.22, 1.24)	0.173
Colonisation with *Pseudomonas aeruginosa*	−0.28 (−1.12, 0.57)	0.520
ABPA	−0.88 (−1.88, 0.12)	0.085
CFRD	0.92 (−0.12, 1.95)	0.084
Pulmonary exacerbation	−0.15 (−0.91, 0.61)	0.695

Risk factors were measured at baseline. ABPA, allergic bronchopulmonary aspergillosis; BMI, body mass index; CFRD, cystic fibrosis related diabetes, FEV_1%pred, forced expiratory volume in 1 second as a percentage of predicted value; FVC%pred, forced vital capacity as a percentage of predicted value; ICS, inhaled corticosteroids; PEx, pulmonary exacerbations; PPI, proton pump inhibitors. *p<0.05

Table S3. Adjusted odds ratios for future pulmonary exacerbation, multivariable model for total cohort (N=545)

Risk factor	**OR (95% CI)**	p-value
Male sex	0.992 (0.729, 1.351)	0.959
Age	0.994 (0.945, 1.046)	0.827
F508 del homozygous	0.903 (0.658, 1.239)	0.527
BMI *z*-score	1.007 (0.841, 1.206)	0.939
FEV_1%pred	0.985 (0.974, 0.996)	0.007*
Pancreatic enzymes	1.252 (0.696, 2.251)	0.453
Prophylactic inhaled antibiotics	1.722 (1.170, 2.533)	0.006*
Prophylactic macrolide	1.346 (0.939, 1.931)	0.107
PPI use	1.565 (1.138, 2.151)	0.006*
Inhaled corticosteroids	1.129 (0.790, 1.613)	0.506
Colonisation with *Pseudomonas aeruginosa*	1.150 (0.751, 1.759)	0.521
ABPA	1.495 (0.882, 2.534)	0.136
CFRD	0.911 (0.504, 1.645)	0.757
Pulmonary exacerbation at baseline	1.901 (1.272, 2.841)	0.002*

Risk factors were measured at baseline. ABPA, allergic bronchopulmonary aspergillosis; BMI, body mass index; CFRD, cystic fibrosis related diabetes, FEV_1%pred, forced expiratory volume in 1 second as a percentage of predicted value; FVC%pred, forced vital capacity as a percentage of predicted value; ICS, inhaled corticosteroids; PEx, pulmonary exacerbations; PPI, proton pump inhibitors. *p<0.05

Table S4. Estimated adjusted effect on mean annual change of FEV_1% predicted, complete case analysis (N=278)

Risk factor	**Effect (95% CI)**	**p-value**
Male sex	0.22 (-0.63, 1.08)	0.606
Age	-0.098 (-0.244, 0.047)	0.183
F508 del homozygous	-0.21 (-1.16, 0.74)	0.664
BMI *z*-score	0.220 (-0.263, 0.703)	0.368
FEV_1%pred	-0.015 (-0.043, 0.014)	0.311
Pancreatic enzymes	-1.17 (-2.68, 0.35)	0.130
Prophylactic inhaled antibiotics	-0.75 (-1.86, 0.37)	0.186
Prophylactic macrolide	0.71 (-0.39, 1.81)	0.201
PPI use	-0.79 (1.67, 0.09)	0.079
Inhaled corticosteroids	0.59 (-0.40, 1.59)	0.240
Colonisation with *Pseudomonas aeruginosa*	-0.36 (1.62, 0.89)	0.566
ABPA	-1.18 (-2.54, 0.18)	0.087
CFRD	0.81 (-0.71, 2.34)	0.291
Pulmonary exacerbation	-0.68 (-1.66, 0.31)	0.178

Risk factors were measured at baseline. ABPA, allergic bronchopulmonary aspergillosis; BMI, body mass index; CFRD, cystic fibrosis related diabetes, FEV_1%pred, forced expiratory volume in 1 second as a percentage of predicted value; FVC%pred, forced vital capacity as a percentage of predicted value; ICS, inhaled corticosteroids; PEx, pulmonary exacerbations; PPI, proton pump inhibitors. *p<0.05

Table S5. Adjusted odds ratios for future pulmonary exacerbation, complete case analysis (N=278)

Risk factor	**OR (95% CI)**	p-value
Male sex	0.877 (0.576, 1.336)	0.541
Age	0.976 (0.912, 1.044)	0.473
F508 del homozygous	0.916 (0.580, 1.447)	0.707
BMI *z*-score	0.964 (0.761, 1.220)	0.760
FEV_1%pred	0.983 (0.969, 0.997)	0.017*
Pancreatic enzymes	1.356 (0.626, 2.936)	0.440
Prophylactic inhaled antibiotics	1.395 (0.820, 2.374)	0.220
Prophylactic macrolide	1.360 (0.820, 2.256)	0.233
PPI use	1.262 (0.825, 1.929)	0.283
Inhaled corticosteroids	1.560 (0.975, 2.495)	0.063
Colonisation with *Pseudomonas aeruginosa*	1.043 (0.569, 1.913)	0.891
ABPA	1.357 (0.696, 2.643)	0.370
CFRD	1.190 (0.583, 2.431)	0.632
Pulmonary exacerbation at baseline	2.664 (1.681, 4.219)	<0.001*

Risk factors were measured at baseline. ABPA, allergic bronchopulmonary aspergillosis; BMI, body mass index; CFRD, cystic fibrosis related diabetes, FEV_1%pred, forced expiratory volume in 1 second as a percentage of predicted value; FVC%pred, forced vital capacity as a percentage of predicted value; ICS, inhaled corticosteroids; PEx, pulmonary exacerbations; PPI, proton pump inhibitors. *p<0.05

References

1. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software 2011; 45(3): 1-67.