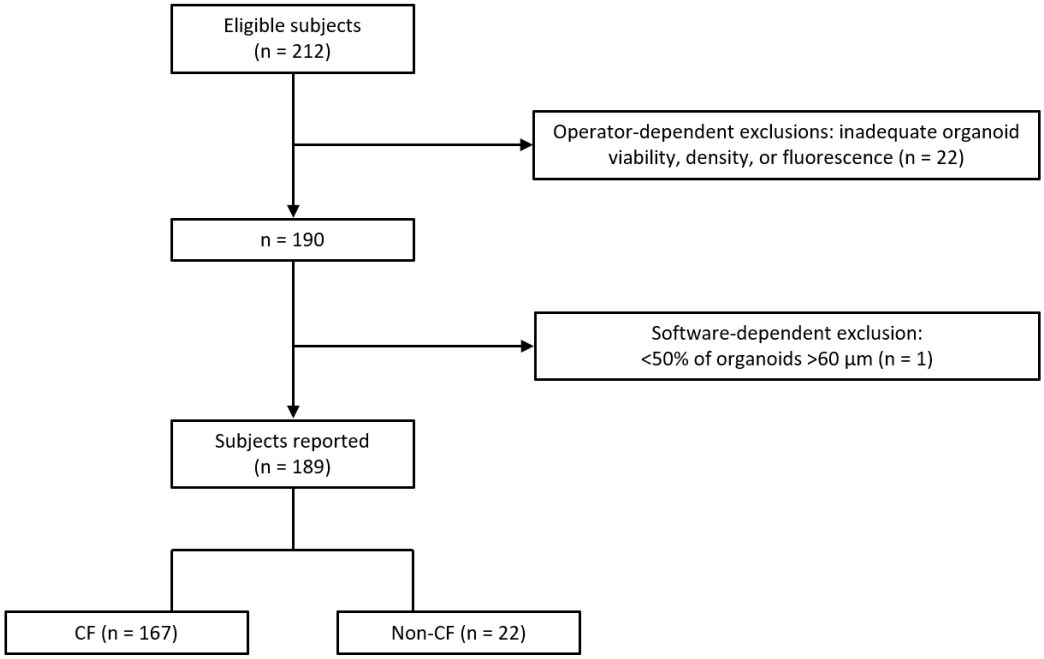
Supplements

Supplement 1: STARD flow diagram



Supplementary figure 1: STARD flow diagram

Supplement 2: Supplementary methods

2.1) Quality criteria for exclusion of organoid cultures

Exclusion criteria for organoid images	
Determined by the operator	**Determined by the software**
Many differentiated or dead structures or debris	Less than 50% of structures >60 µm (organoids should be large enough to show morphology typical to either CF or non-CF)
Inadequate plating density: too many (eg. overlapping) or too few organoids	<500 or >3000 organoids (defined as structures >40 µm) in 32 wells; <125 or >750 organoids in 8 wells
Inadequate fluorescence distribution (organoids not clearly delineated, background signal too high)	

Supplementary table 1: quality criteria for exclusion of organoid cultures

2.2) Plating of organoids and image acquisition

Organoids were plated in 32 wells of a 96 well plate, left to incubate overnight (16-24 hours) without CFTR modulators or forskolin, and stained with calcein green 0.6 µM. The plate was then transferred to a LMS800 Zeiss confocal microscope with an automated stage and integrated incubator. The x/y position and the best focus (z position) of the organoids in each well were manually defined. Images were then automatically acquired in a unidirectional way using a 5x objective with a resolution of 1024x1024 pixels (pixel size 2.5x2.5 µm) and 16 bits depth. Zen blue software (v2.6 from Zeiss) was used. The 32 images per subject were exported as TIFF files and analysed in the NIS-Elements Advanced Research Analysis Imaging Software (v.5.02.00) according to the table below.

2.3) Semi-automated image analysis protocol

A semi-automated protocol was used for image analysis, in which the number of organoids (defined as structures ≥40 µm) was counted, as well as the number of organoids which had grown enough to show the typical CF or non-CF morphology (defined as structures ≥60 µm). The exclusion of organoids ≥40 and <60 µm was based on separate analysis of this subset, showing no significant morphological differences for these small structures between CF and non-CF subjects (round and dense in both cases). Two erosion steps were performed on all included structures to analyse the whole organoid and the central part separately. Two indices were calculated to quantify organoid morphology (figure 2).

Semi-automated protocol for image analysis in NIS-Elements Advanced Research Analysis Imaging Software (v.5.02.00)	
Procedure	Rationale
Make one Network Data (.ND) file of all 32 pictures for each subject	Enables simultaneous analysis of all 32 pictures per subject
Recalibrate images so 1 pixel is 2.5x2.5 µm ("recalibrate document" function)	Enables erosions in µm instead of pixels
Delineate structures with a lower intensity threshold of 4500 and an upper threshold of 65535 ("define threshold" function); "smooth" and "clean" functions off; “fill holes” function on; "separate" function on x3	Delineates fluorescent structures, including the lumen if present
Select all structures ≥40 µm and count ("size" parameter)	Counts the total number of organoids in the 32 wells, excludes small debris
Select all structures ≥60 µm and count ("size" parameter)	Counts the organoids large enough to show morphology typical to either CF or non-CF
Remove all structures touching the borders of the picture ("remove objects touching border" function)	Removes organoids that are not completely visible, as their morphology cannot be accurately quantified
Erode 1 pixel (=2.5 µm) from the border of each ≥60 µm structure ("erode" function)	Removes the “halo” of background fluorescence surrounding the organoids
Measure the mean intensity of each structure	Measures the mean fluorescence of the whole organoids
Erode 10 pixels (=25 µm) from the border of each original ≥60 µm structure ("erode" function)	In non-CF organoids, this erodes the cellular border and leaves only the lumen
	In CF organoids, this erodes the outer part of the organoid, which has roughly the same intensity as the inner part that remains
Measure the mean intensity of each structure	Measures the mean fluorescence of the central part of the organoids
Measure the circularity of each structure	Measures the mean circularity of the organoids

Supplementary table 2: semi-automated protocol for image analysis

2.4) Statistical methods and additional analyses

Linear discriminant analysis and quadratic discriminant analysis were used to discriminate between subjects with and without CF using IR and CI indexes. The area under the ROC curve (AUC) was reported. A bootstrap approach was used to obtain a 95% confidence interval for this AUC and to estimate the degree of overoptimism in its quantification. Indeed, the observed AUC overestimates the future performance since the same data were used to construct and evaluate the model. Given that the data are sparse at the intersection of both groups, a parametric resampling procedure was preferred to a non-parametric version to quantify the degree of overoptimism in the estimation of the AUC.

Step 1: sample a mean vector and a covariance matrix (for IR and CI) from the observed mean and covariance.

Step 2: draw N0 controls and N1 cases, with N0 and N1 the observed number of controls and cases, respectively. They form one bootstrap sample.

Step 3: calculate the AUC in the bootstrap sample and the AUC in the original dataset based on the model in the bootstrap sample.

Repeat step 1 to step 3 1000 times.

Step 4: calculate the mean AUC of the 1000 bootstrap samples and the mean AUC of the 1000 evaluations in the original dataset. The difference is the estimate of overoptimism. Subtract this estimate from the observed AUC to get the optimism-corrected AUC.

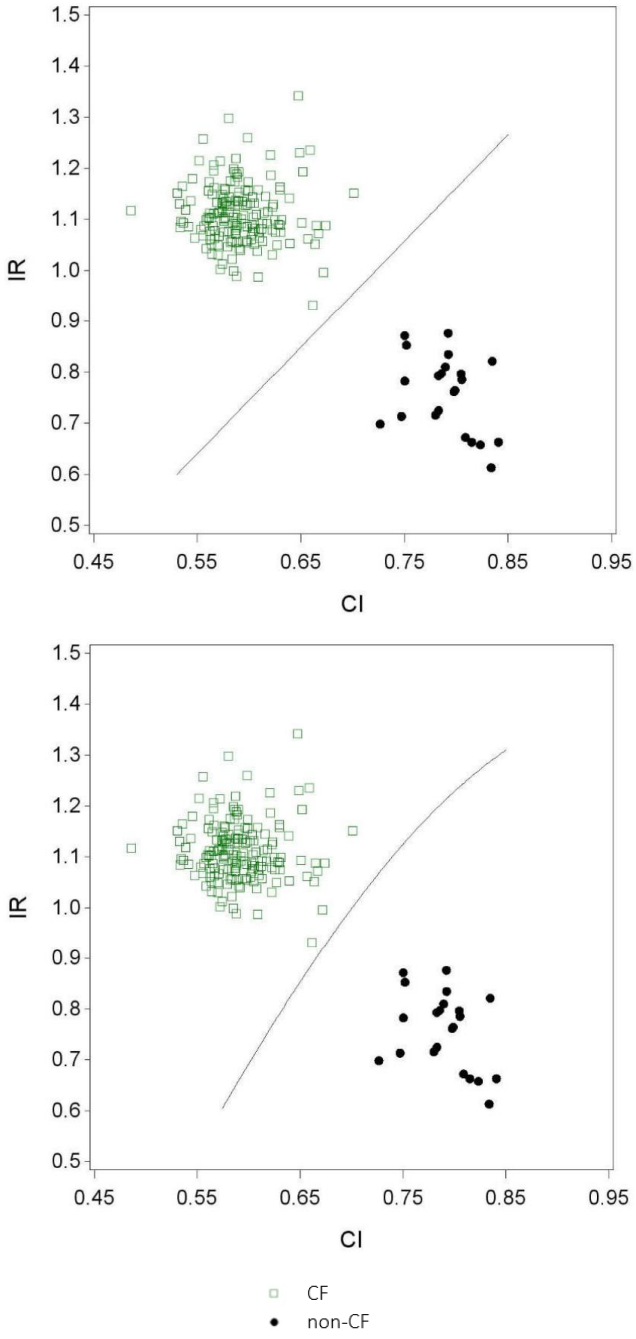
Perfect discrimination (AUC=1) between CF and non-CF using IR and CI was obtained with linear as well as with quadratic discriminant analysis. Perfect discrimination was also obtained in all bootstrap samples (hence, the lower limit of the confidence interval also equals 1). The AUC corrected for overoptimism equalled 0.9879132 and 0.9998538, respectively.

These analyses were performed using SAS software, version 9.4 of the SAS System for Windows.

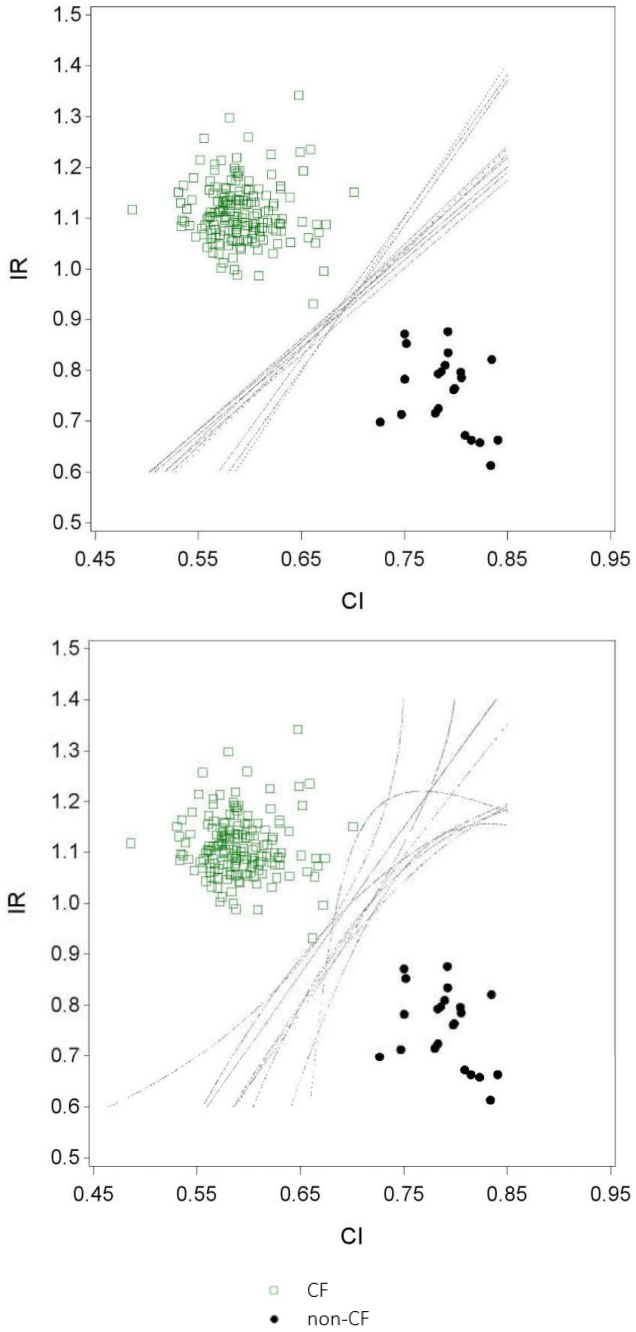
Supplementary figure 2: Results for linear (upper panel) and quadratic (lower panel) discriminant analysis.

Discrimination lines were obtained with linear and quadratic discriminant analysis, respectively. These lines correspond to combinations of values for which the posterior probability to be a case equals 0.50 (using an equal prior probability for both classes).

CI: circularity index; IR: intensity ratio

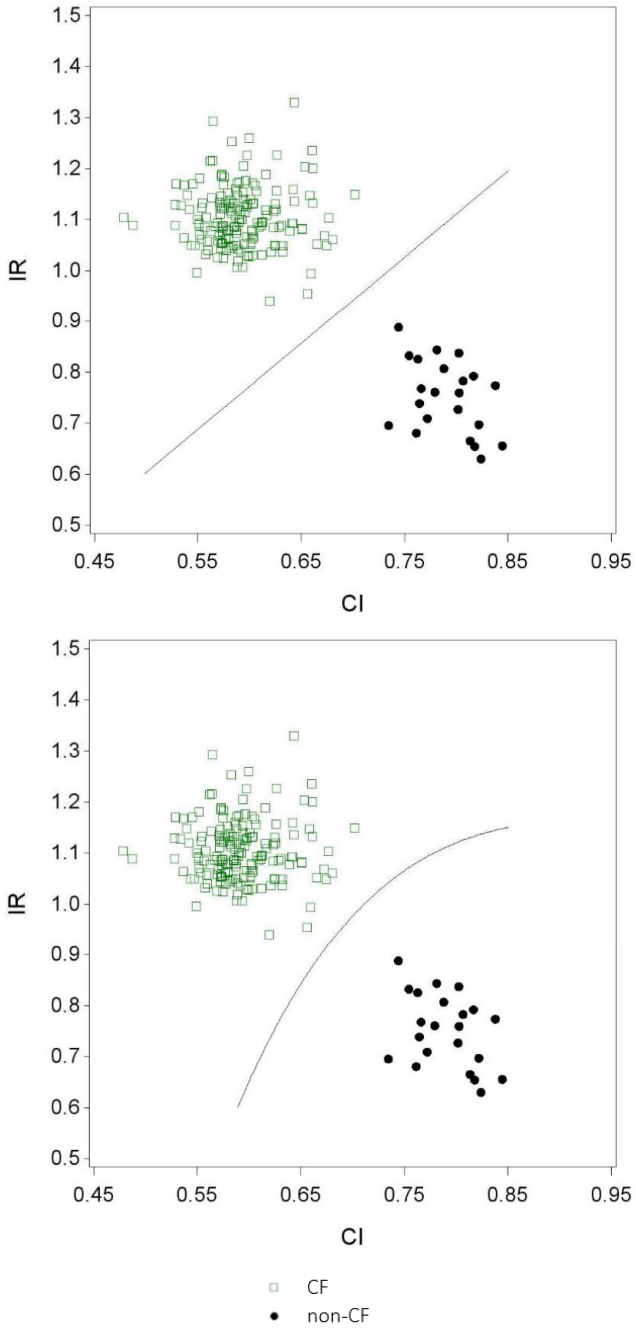


Supplementary figure 3: Discrimination lines obtained in 10 bootstrap samples for linear (upper panel) and quadratic (lower panel) discriminant analysis.
These discrimination lines give an impression of the uncertainty of the discrimination line by presenting the result in 10 bootstrap samples.
CI: circularity index; IR: intensity ratio

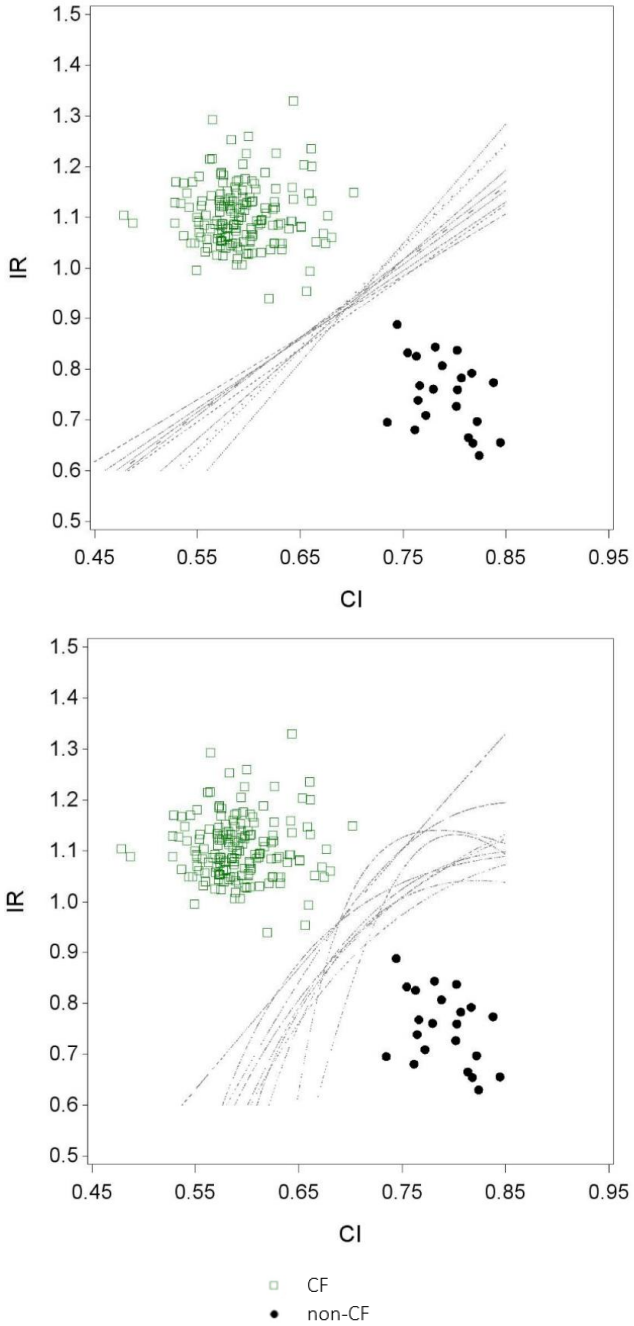


Results for data using 8 wells at random per subject and using the same statistical methodology were comparable to the results using 32 wells. Again, perfect discrimination was obtained with the linear as well as with the quadratic discriminant analysis.

Supplementary figure 4: results for linear (upper panel) and quadratic (lower panel) discriminant analysis, based on 8 wells data.
CI: circularity index; IR: intensity ratio



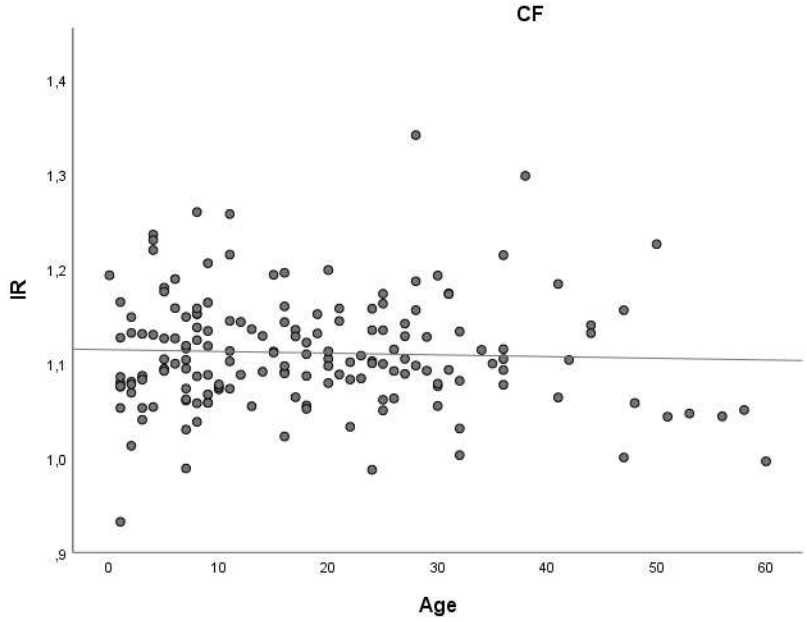
Supplementary figure 5: Discrimination lines obtained in 10 bootstrap samples, based on 8 wells data. Results for linear (upper panel) and quadratic (lower panel) discriminant analysis.
CI: circularity index; IR: intensity ratio



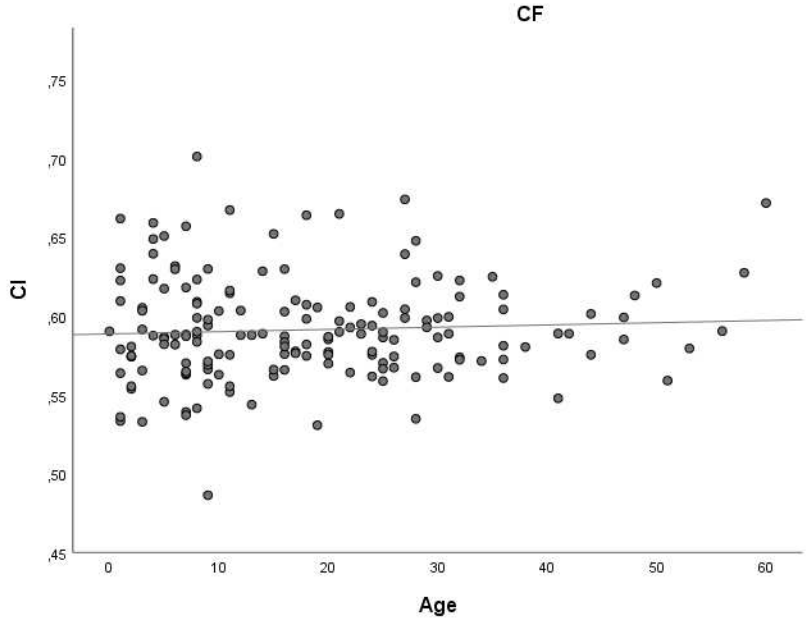
Supplement 3: Supplementary statistical tests for subgroup analysis

Additional statistical test were performed for subgroup analysis. Pearson's correlation coefficients were used to assess correlations between ROMA indexes and age, and Mann-Whitney and Fisher exact tests for between-group comparisons in IBM SPSS Statistics Version 26.

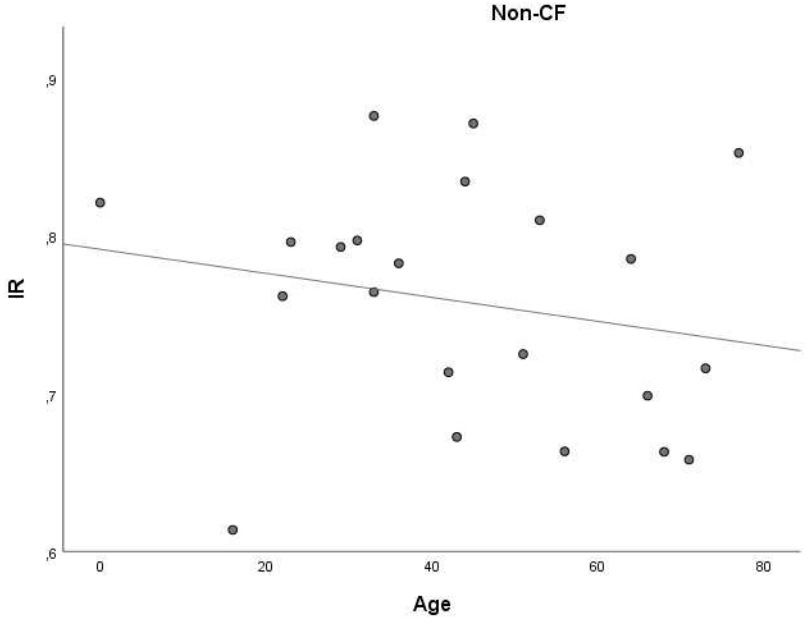
Supplementary figure 6: scatter plot with fit line for intensity ratio (IR) by age in CF subjects (Pearson's correlation coefficient *r*=-0.042 (p=0.592))



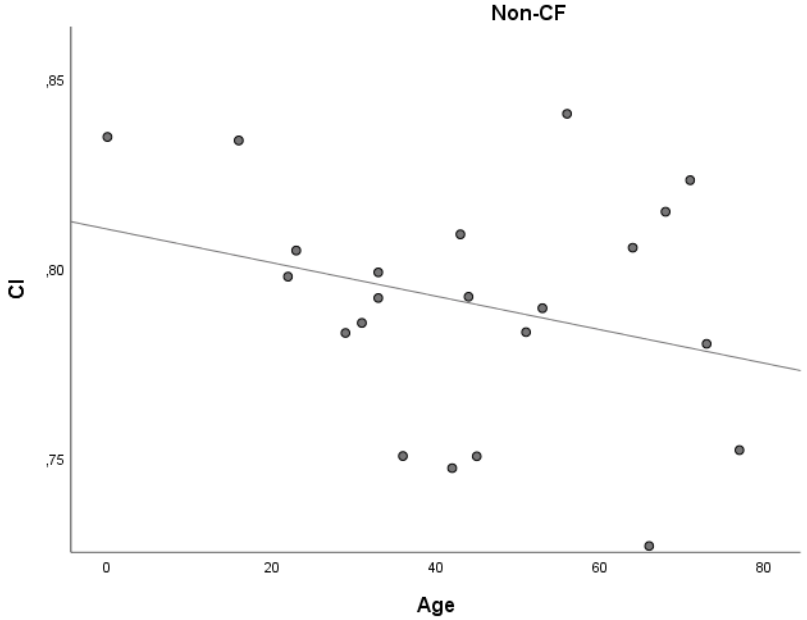
Supplementary figure 7: scatter plot with fit line for circularity index (CI) by age in CF subjects (Pearson's correlation coefficient *r*=0.060 (p=0.440))



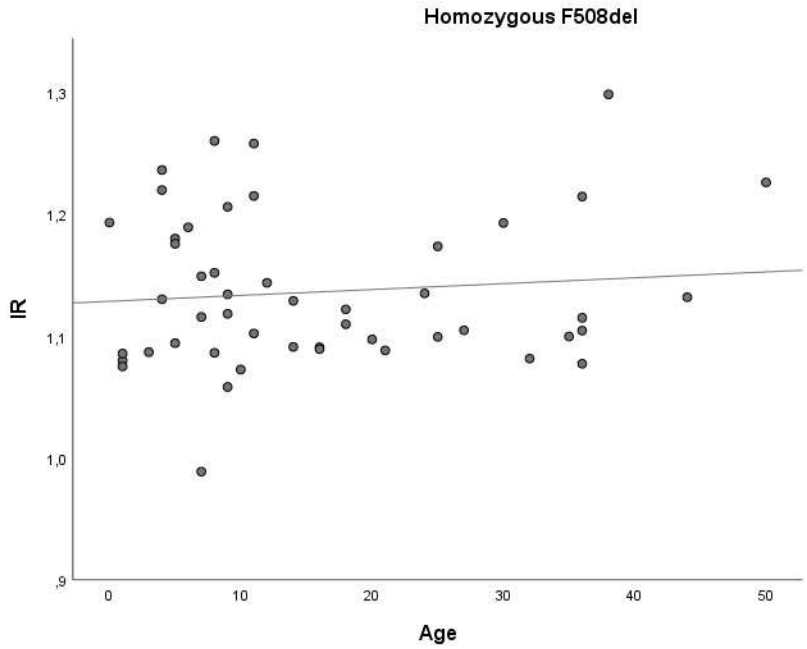
Supplementary figure 8: scatter plot with fit line for intensity ratio (IR) by age in non-CF subjects (Pearson's correlation coefficient *r*=-0.207 (p=0.356))



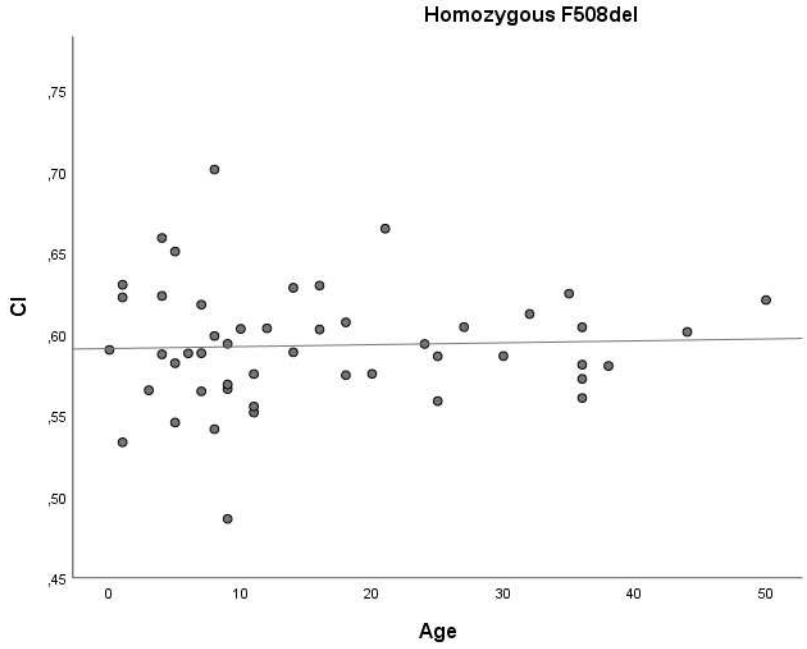
Supplementary figure 9: scatter plot with fit line for circularity index (CI) by age in non-CF subjects (Pearson's correlation coefficient *r*=-0.294 (p=0.185))



Supplementary figure 10: scatter plot with fit line for intensity ratio (IR) by age in homozygous F508del subjects (Pearson's correlation coefficient *r*=0.100 (p=0.494))



Supplementary figure 11: scatter plot with fit line for circularity index (CI) by age in homozygous F508del subjects (Pearson's correlation coefficient *r*=0.042 (p=0.777))



Supplementary table 3: between-group differences for subjects with pancreatic sufficiency and pancreatic insufficiency

	PS	PI	p value
n	28	137	
IR	1.084 (0.987-1.165)	1.117 (0.932-1.342)	0.007
CI	0.597 (0.531-0.672)	0.591 (0.486-0.701)	0.419
SCC low (<87 mmol/L) or high (≥87 mmol/L)	16 low (59%) 11 high (41%)	24 low (18%) 111 high (82%)	<0.001

n or mean and range

PI: pancreatic insufficient; PS: pancreatic sufficient; IR: intensity ratio; CI: circularity index; SCC: sweat chloride concentration

Supplementary table 4: between-group differences for subjects with low and high sweat chloride

	SCC <87 mmol/l	SCC ≥87 mmol/l	p value
n	41	123	
IR	1.11 (0.99-1.22)	1.11 (0.93-1.34)	0.763
CI	0.59 (0.49-0.66)	0.59 (0.53-0.70)	0.602
Pancreatic status (n = 162)	16 PS (40%) 24 PI (60%)	11 PS (9%) 111 PI ((91%)	<0.001

n or mean and range

SCC: sweat chloride concentration; IR: intensity ratio; CI: circularity index; PS: pancreatic sufficient; PI: pancreatic insufficient

Supplementary table 5: subjects ordered by genotype with age, sweat chloride concentration, pancreatic status, and ROMA indexes

Genotype	n	Range age	Range sweat chloride concentration	PS/PI	Range CI	Range IR
F508del / F508del	49	0 - 50	73,5 - 134	0 / 49	0,49 - 0,70	0,99 - 1,30
F508del / G542X	11	4 - 31	79 - 112,2	0 / 11	0,53 - 0,65	1,06 - 1,34
F508del / 3272-26A->G	8	3 - 47	65,1 - 102	5 / 3	0,53 - 0,66	1,00 - 1,16
F508del / L927P	7	2 - 24	74,1 - 160	0 / 7	0,53 - 0,62	1,05 - 1,20
F508del / N1303K	7	1 - 9	81,1 - 121	0 / 7	0,56 - 0,66	0,93 - 1,10
F508del / 1717-1G->A	4	16 - 25	74,1 - 96	0 / 4	0,57 - 0,61	1,10 - 1,16
F508del / A455E	4	6 - 60	71 - 91,1	3 / 1	0,57 - 0,67	1,00 - 1,16
F508del / E92K	3	2 - 27	117 - 137	1 / 1	0,54 - 0,64	1,08 - 1,14
F508del / 2789+5G->A	3	1 - 56	74 - 122	2 / 1	0,54 - 0,6	1,04 - 1,16
F508del / 3849+10kbC->T	3	8 - 29	36 - 113	3 / 0	0,58 - 0,6	1,06 - 1,13
1717-1G->A / L927P	2	6 - 24	97,1 - 107	0 / 2	0,56 - 0,58	1,10 - 1,13
F508del / G970R	2	1 - 20	83 - 87,1	0 / 2	0,58 - 0,61	1,08 - 1,10
F508del / I507del	2	18 - 20	87 - 114	0 / 2	0,59 - 0,60	1,08 - 1,09
F508del / R553X	2	3 - 16	88,1 - 94,1	0 / 2	0,58 - 0,59	1,02 - 1,04
F508del / S1255P	2	25 - 31	92 - 97	0 / 2	0,56 - 0,59	1,14 - 1,17
F508del / Y1092X	2	7 - 22	89,1 - 93	0 / 2	0,56 - 0,61	1,06 - 1,08
N1303K / delEx19	2	9 - 22	85,1 - 87,1	0 / 2	0,56 - 0,63	1,03 - 1,16
N1303K / N1303K	2	10 - 31	82,1 - 94,1	0 / 2	0,56 - 0,59	1,08 - 1,09
F508del / R334W	2	22 - 26	94 - 98	1 / 1	0,57 - 0,59	1,06 - 1,10
W1282X / I1234V	2	25 - 30	79 - 115	1 / 1	0,57 - 0,63	1,05 - 1,08
F508del / I336K	2	21 - 30	83 - 85	2 / 0	0,60 - 0,60	1,06 - 1,16
F508del / 182delT	1	9	59	0 / 0	0,56	1,09
1811+1.6kbA->G / 1811+1.6kbA->G	1	32	116,8	0 / 1	0,57	1,00
2143delT / G542X	1	1	110	0 / 1	0,58	1,13
2143delT / N1303K	1	9	84,89	0 / 1	0,60	1,06
3272-26A->G / L165S	1	44	99	0 / 1	0,58	1,14
394delTT / L927P	1	11	89,1	0 / 1	0,61	1,11
405+1G->A / 3121-1G->A	1	8	102	0 / 1	0,62	1,12
405+1G->A / 405+1G->A	1	27	105	0 / 1	0,67	1,09
711+1G->T / c.1819_1902del	1	13	102	0 / 1	0,59	1,05
c.1329_1350del / c.c.1329_1350del	1	2	85,1	0 / 1	0,58	1,08
F508del / 1078delT	1	15	86,1	0 / 1	0,65	1,19
F508del / 1259insA	1	17	116	0 / 1	0,58	1,14
F508del / 1898+1G->C	1	41	113	0 / 1	0,59	1,18
F508del / 394delTT	1	28	77	0 / 1	0,62	1,19
F508del / delEx19,20,21	1	10	106	0 / 1	0,58	1,08
F508del / E60X	1	34	98	0 / 1	0,57	1,11
F508del / G551D	1	32	98	0 / 1	0,62	1,03

F508del / L227R	1	13	93	0 / 1	0,54	1,14
F508del / M1101K	1	20	108	0 / 1	0,57	1,11
F508del / R352Q	1	41	59,9	0 / 1	0,55	1,06
F508del / S1251N	1	53	91	0 / 1	0,58	1,05
F508del / W1282X	1	36	106	0 / 1	0,61	1,09
F508del / W401X	1	42	121	0 / 1	0,59	1,10
G542X / G970R	1	26	85,1	0 / 1	0,57	1,11
I507del / L927P	1	7	96	0 / 1	0,54	1,09
L227R / L227R	1	1	114,5	0 / 1	0,56	1,05
L927P / L927P	1	11	89,1	0 / 1	0,67	1,07
N1303K / 1717-1G->A	1	7	111	0 / 1	0,59	1,07
N1303K / 4010del4	1	23	125	0 / 1	0,59	1,08
N1303K / Q359K/T360K	1	17	117	0 / 1	0,61	1,06
Q493X / W401X	1	5	82,1	0 / 1	0,59	1,13
R1158X / R1162X	1	16	102	0 / 1	0,58	1,16
R553X / A455E	1	7	83,1	0 / 1	0,56	1,10
W1282X / delEx2-3	1	28	130	0 / 1	0,56	1,16
W1282X / G85E	1	30	113,8	0 / 1	0,57	1,08
W1282X / W1089X	1	8	118	0 / 1	0,59	1,15
1717-1G->A / 2789+5G->A	1	7	79,1	1 / 0	0,66	1,06
2183AA->G / R347H	1	26	68,1	1 / 0	0,59	1,09
3272-26A->G / 3659delC	1	16	91,1	1 / 0	0,58	1,14
A455E / delEx1 and promoter	1	48	98	1 / 0	0,61	1,06
F508del / L138ins	1	2	108	1 / 0	0,57	1,13
F508del / R117C	1	4	46,5	1 / 0	0,64	1,05
G542X / 2789+5G->A	1	12	98,1	1 / 0	0,59	1,09
I507del / E60K	1	24	73	1 / 0	0,61	0,99
N1303K / I336K	1	51	85	1 / 0	0,56	1,04
W1282X / 3849+10kbC->T	1	58	124	1 / 0	0,63	1,05

PS: pancreatic sufficient; PI: pancreatic insufficient; CI: circularity index; IR: intensity ratio