



# Stochastic Tracking of Infection in a CF Lung

Sara Zarei<sup>1\*</sup>, Ali Mirtar<sup>2</sup>, Forest Rohwer<sup>3</sup>, Peter Salamon<sup>4</sup>

**1** Computational Science Research Center, San Diego State University, San Diego, California, United States of America, **2** Electrical and Computer Eng. Dep/University of California San Diego, San Diego, California, United States of America, **3** Department of Biology, San Diego State University, San Diego, California, United States of America, **4** Department of Mathematics and Statistics, San Diego State University, San Diego, California, United States of America

## Abstract

Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scan are the two ubiquitous imaging sources that physicians use to diagnose patients with Cystic Fibrosis (CF) or any other Chronic Obstructive Pulmonary Disease (COPD). Unfortunately the cost constraints limit the frequent usage of these medical imaging procedures. In addition, even though both CT scan and MRI provide mesoscopic details of a lung, in order to obtain microscopic information a very high resolution is required. Neither MRI nor CT scans provide micro level information about the location of infection in a binary tree structure the binary tree structure of the human lung. In this paper we present an algorithm that enhances the current imaging results by providing estimated micro level information concerning the location of the infection. The estimate is based on a calculation of the distribution of possible mucus blockages consistent with available information using an offline Metropolis-Hastings algorithm in combination with a real-time interpolation scheme. When supplemented with growth rates for the pockets of mucus, the algorithm can also be used to estimate how lung functionality as manifested in spirometric tests will change in patients with CF or COPD.

**Citation:** Zarei S, Mirtar A, Rohwer F, Salamon P (2014) Stochastic Tracking of Infection in a CF Lung. *PLoS ONE* 9(10): e111245. doi:10.1371/journal.pone.0111245

**Editor:** Wayne Iwan Lee Davies, University of Western Australia, Australia

**Received:** June 16, 2014; **Accepted:** September 22, 2014; **Published:** October 31, 2014

**Copyright:** © 2014 Zarei et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability:** The authors confirm that all data underlying the findings are fully available without restriction. Data are all contained within the paper.

**Funding:** This material is based upon work supported by the National Institutes of Health under Grant no. 56586B. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* Email: sara.zarei@alumni.ucsd.edu

## Introduction

Patients with chronic obstructive pulmonary disease (COPD) or cystic fibrosis (CF) have chronic lung inflammation which causes airflow limitation and the scarring of lung tissues. Their airways are generally inflamed and produce excess amounts of mucus that impair the flow of air into and out of their lungs.

Our hypothesis is that the scarring, and ultimate remodeling of a CF lung is mostly due to the contact between the lung lining and the mucus. Inflammatory cytokines induce scars in lung tissue [1]. This contact between mucous biofilm and lung tissue facilitates virulent microbes that also play a role in remodeling a CF lung. In comparison with a normal lung, the airway fluids of CF patients contain large amounts of neutrophils as a result of the inflammatory response [2]. Thus our hypothesis is that mucus accumulation is primarily responsible for damage to the lung and therefore tracking its propagation is a crucial task for better diagnosis and treatment.

Spirometry, the measurement of a patient's breathing, is the standard clinical tool for monitoring lung disease. The two most common spirometric indicators are the forced expiratory volume in one second, FEV<sub>1</sub>, and the forced vital capacity, FVC. FEV<sub>1</sub> measures the volume of air that can forcibly be blown out in one second while FVC measures the volume of air that can forcibly be blown out after a full inspiration maneuver. Although these tests provide a global measure of airflow obstruction and restriction, they do not give detailed information about the location of mucus blockage. Additional information is available from the repeated imaging of CF patients' lungs. This is usually in the form of chest

x-ray, computed tomography (CT) or magnetic resonance imaging (MRI). Despite the fact that CT is an optimal morphological assessment of CF lung changes, the associated exposure to ionizing radiation is a serious obstacle. Therefore, MRI might be the appropriate method for imaging CF patient's lungs [3,4].

MRI was first introduced as an alternative imaging tool for patients with CF in 1987 [5]. There are numerous methods for the analysis of chest MRI. Theilmann et al. proposed a new MRI imaging method that can spatially locate the pockets of infection and measure the amount of mucus located within each pocket. This information is obtained with a resolution of approximately 1.0 cm<sup>3</sup> [6]. This resolution is equivalent to the last 10 generations of airways combined. Despite the accuracy level of MRI images, they do not contain any micro-level information on smaller airways. Hence the precise location of infection cannot be obtained from the MRI data.

There have been research studies on mesoscopic modeling of Cystic Fibrosis [7]. Such studies however did not address the micro level information about mucus propagation through the airways. The goal of this research is to provide the clinician with an algorithm that can track the location and propagation of mucus in a CF patient's lungs. Tracking the growth, or shrinkage of these pockets can be correlated to the efficacy of different treatment regimens on each pocket. Assuming further progress in metagenomic and transcriptomic analyses of sputum samples, tracking may allow also correlating the growth or shrinkage of these pockets and the composition of the local microbial community structure within the pocket.

The central tool introduced in the present article for tracking pockets of inflammation is based on using our airflow model [8] in reverse, i.e., as the main ingredient in an inverse problem. Our airflow model can calculate the flow and the total resistance corresponding to a given distribution of mucus obstructions in respiratory airways. The solution of the inverse problem is achieved by randomly sampling the many possible mucus configurations consistent with the current information regarding a patient. Many micro-scale distributions match any observed MRI and spirometric data. The fortuitous finding in the calculations described below is that a large majority of the possible distributions fall within narrow ranges of certain parameters such as which generations of airways contain how much mucus. Can we be certain of these most likely locations being the actual mucus distribution? The fact that most of the distributions consistent with the spirometric and MRI data have these features makes them a good bet while not giving us certainty that any *one* configuration is of this form. In fact, however, repeated MRI measurements separated by small challenges such as coughing or even just taking a few deep breaths reveal that mucus inside the lungs of CF patients show small but discernible movements in response to such challenges. Given this dynamic picture, the possibility that all of the configurations of mucus would avoid the configurations that can be realized the largest number of ways is extremely unlikely. We describe below how to construct such distributions that best represent the state of airways corresponding to the patients lung functionality test values (FEV<sub>1</sub> and FVC) and any available MRI or CT data. We use a maximum entropy approach to construct such estimates by sampling all consistent distributions and choosing the one that can be realized the largest number of ways. We find that the predicted microscopic distribution of mucus is generally sharply peaked, allowing us to estimate the distribution of mucus as the one that appears the most frequently in our simulation. The entire computation process can take a long time and would require enormous computational space to run the simulation for the entire 2<sup>23</sup> airways in a lung. However after running the simulation once for a patient, we can store most of the data so the next simulations only take a few minutes to complete.

Given the micro-scale distribution obtained from our inverse problem, we can use our physiological model [9] to predict the growth and propagation of this distribution, thereby predicting the progression of the disease. This second model requires a rate for the growth of the mucus volume at each location. Currently only one overall average growth rate is available, a growth rate that was estimated based on a forty year CF population average [9]. Predictions of this model with the average growth rate can nonetheless be compared to actual observed disease progression between exacerbations. For predictions during exacerbations, a database of treatment and community specific growth rates are needed and in-principle available from multiple MRI and metagenomic/transcriptomic analyses. Two such datasets allow the extraction of pocket specific growth rates and corresponding linear extrapolation for the mucus volume in each pocket. This can at least lengthen the times between imaging sessions.

## Methods

### Inverse Problem Using Metropolis-Hastings Algorithm

Our algorithm uses CF patients' spirometry test values of FVC, FEV<sub>1</sub> and any available mucus distribution data from CT scans or MRI as its input. It then identifies the distribution of mucus obstructed bronchioles throughout different airway generations. The model assumes the lung airways to be binary branching trees [10] extending over 23 generations from the bronchus down to the

alveoli. It further assumes a fractal structure [11–13] for the parameters of the binary airway tree.

Considering the fact that there are 2<sup>23</sup> bronchioles in a human lung, there are astronomically many possible configurations for a certain amount of mucus distributed in an airway tree, even given the mucus in each voxel. In fact many of these configurations will result in the same FEV<sub>1</sub> and FVC values. Since exhaustive sampling of these configurations is impossible, we use the Metropolis-Hastings Markov chain Monte Carlo algorithm to sample a few hundred million and base our estimates on such a sample. The goal of this section is to introduce this method.

Our algorithm proceeds from an assumed amount of mucus to be distributed into the 2<sup>23</sup> airways. It starts from a random distribution and obtains an unbiased sample of configurations satisfying certain requirements. These requirements are implemented as soft constraints via an energy function

$$E = (\alpha)(\text{Calculated FEV}_1 - \text{Patient FEV}_1)^2 + (1 - \alpha)(\text{Calculated FVC} - \text{Patient FVC})^2 \quad (1)$$

where  $\alpha$  and  $(1 - \alpha)$  are weight factors. In our algorithm we set  $\alpha = 0.5$ . The energy function is used as a way to force the sampling to stay near values with low energy, i.e. configurations with approximately correct spirometric readings.

We then examine successive samples by performing a random walk on the space of mucus configurations. Each move in our random walk reallocates the location of some of the mucus. The resulting distribution of mucus configurations turns out to be sharply peaked in certain natural parameters, a fact exploited by our algorithm.

The Metropolis algorithm is a widely used procedure for sampling a sequence from a specified distribution on a large finite set. It describes equilibrium for systems whose configurations have probability proportional to the Boltzmann factor ( $e^{-E/T}$ ). This is a weighting factor which determines the relative probability that the system will be found in a particular configuration at energy  $E$  when the temperature of the environment is  $T$  [14]. We used a constant temperature,  $T = 1$ , making the probability of a configuration with energy  $E$  proportional to  $\exp(-E)$ .

The following steps describe the Metropolis algorithm [15].

- Initiate the sampling from an arbitrary configuration  $A$ , with known energy  $E_A$ .
- Define a new neighbor configuration  $B$  from configuration  $A$ .
- Calculate the new configuration energy,  $E_B$ . This trial move is then either accepted or rejected according to the following simple probabilistic rule.
- If  $E_B < E_A$  we accept the new configuration.
- if  $E_B > E_A$ , we may accept configuration  $B$  with the following probability.

$$p = e^{-(E_B - E_A)/T} \quad (2)$$

- Repeat until sufficient number of configurations have been collected.

To apply the Metropolis algorithm to our model, we need to define neighboring configurations. Our definition moves a certain amount of mucus between a few chosen bronchioles. This keeps the energy re-evaluation step computationally cheap and, provided we allow such rearrangement of all types of airways, we

avoid the problems created by local minima. This assures that the system settles into and stays near the lowest energy configurations as the simulation proceeds [16].

## Two-Dimensional Probability Density Function Estimation (PDFE-2D) of Mucus Obstructions

Imaging provides us with the data that corresponds to the spatial location of infection pockets and the amount of mucus within each pocket. Both MRI and CT images can have various resolutions and our approach is scalable to any resolution. For concreteness below we work with a resolution of  $1 \text{ cm}^3$  and refer to this smallest volume as a voxel - a volume element.

Since on average the total lung capacity of an adult human is about 6 liters [17], an image will have about 6000 voxels. Using the binary tree structure of the lung airway, we set these voxels at the end of the 13th generation to approximately match the number of these elements to the number of subtrees that remain. Since  $2^{13} = 8192 \approx 6000$ , we can identify the mucus in the voxel with the mucus in the subtree of  $2^{23-13} = 2^{10}$  bronchioles terminating in the alveoli. Thus each voxel in our lung model represents a binary tree structure that has a total of 10 bifurcations, from generation 13 to 23, with a known total amount of mucus obtained from imaging data.

Using a three-dimensional model of the human airway tree that was developed by [18], we are able to map each MRI or CT scan voxel into our lung model's voxels with their corresponding mucus content. The main task of our algorithm is to locate the infection in the airway tree structure of a CF lung. To achieve this goal, we use two summary features for each voxel: (1) the percent of alveoli that are accessible, i.e., not totally blocked and (2) the total resistance to flow from the alveoli to the 13th generation bronchiole assigned to the voxel. These two microscopic features correspond loosely to the spirometric indicators FVC and FEV<sub>1</sub>, respectively.

In order for us to calculate each voxel's resistance and the percent of accessible alveoli (for simplicity, we refer to this as the AA%), we first have to define how mucus is distributed in a voxel's binary tree structure. There is an astronomical number of configurations for filling a binary tree of 10 generations with a given amount of mucus. Approximately, it is given by  $(2^{10})^{M/m}$  where  $M$  is the total mucus volume and  $m$  is the volume of the smallest bronchiole. Thus even at this level, we used Metropolis-Hastings to sample many mucus configurations, recording the values of AA% and resistance for each configuration. We set the amount of mucus within each voxel and calculate the distribution of AA% and resistance: our PDFE-2D distribution.

Hence if a voxel contains 5% mucus, the first step is randomly filling up the voxel's airway tree with the specified mucus amount. Then using the Metropolis algorithm, at each state we move only a fraction of the mucus within a bronchiole that is equivalent to the smallest bronchiole's volume in a lung airway tree. We refer to this as the "unit volume". Once a unit volume is moved to a different location, the corresponding voxel's parameters (AA% and resistance) are recalculated.

In order to expedite the computational process; Dulcinea computing clusters from the Computational Science Research Center at San Diego State University were used for collecting almost 54 million samples. The Dulcinea computing clusters contains 12 workstations each with Dual-Quad Xeon central processing unit (CPU) (E5520 2.27GHz) and Dual Tesla graphic processing unit (GPU) (M1060) which provides the total of 96 CPU cores. The cluster system utilizes 3GB of memory per CPU core for nodes 1 to 10 and utilizes 6GB of memory per CPU core

for nodes 11 and 12. After obtaining these samples the probability distributions for different amount of mucus are calculated. Figure 1A to Figure 1D illustrate the probability density function for (5%, 30%, 65%, and 75%) mucus respectively. As shown in Figure 1A, when there is only 5% mucus in a voxel, the most likely configuration has 86% of its alveoli accessible and the voxel's resistance increases by a factor of almost 1.1. When the mucus level reaches almost 30% there are only 27% accessible alveoli and the voxel resistance is almost 2.2 times a healthy voxel with no mucus. On the other hand in Figure 1C and 1D the number of accessible alveoli value approaches zero while the resistance value reaches infinity. This refers to a case that a voxel is almost completely filled with mucus to an extent that no more air can pass through and therefore blocks all the corresponding alveoli at the end of the branching tree.

Figure 2 displays the maximum likelihood combinations of AA% and resistance ratio for different amount of mucus in a voxel. As the mucus reaches almost 60% of the available airway volume in the voxel, there is no remaining access to the alveoli and as a result there is no gas exchange taking place in that particular part of the airway tree. After collecting these distributions, the model can initiate the prediction steps as well as providing the microlevel information about the location of obstructed bronchioles. We will discuss each outcome in the next two sections. Please note that all data underlying the findings of this section have been discussed in the manuscript. Other than the massive computing power needed to produce the findings in Figure 1 and Figure 2, all the relevant data have been shared.

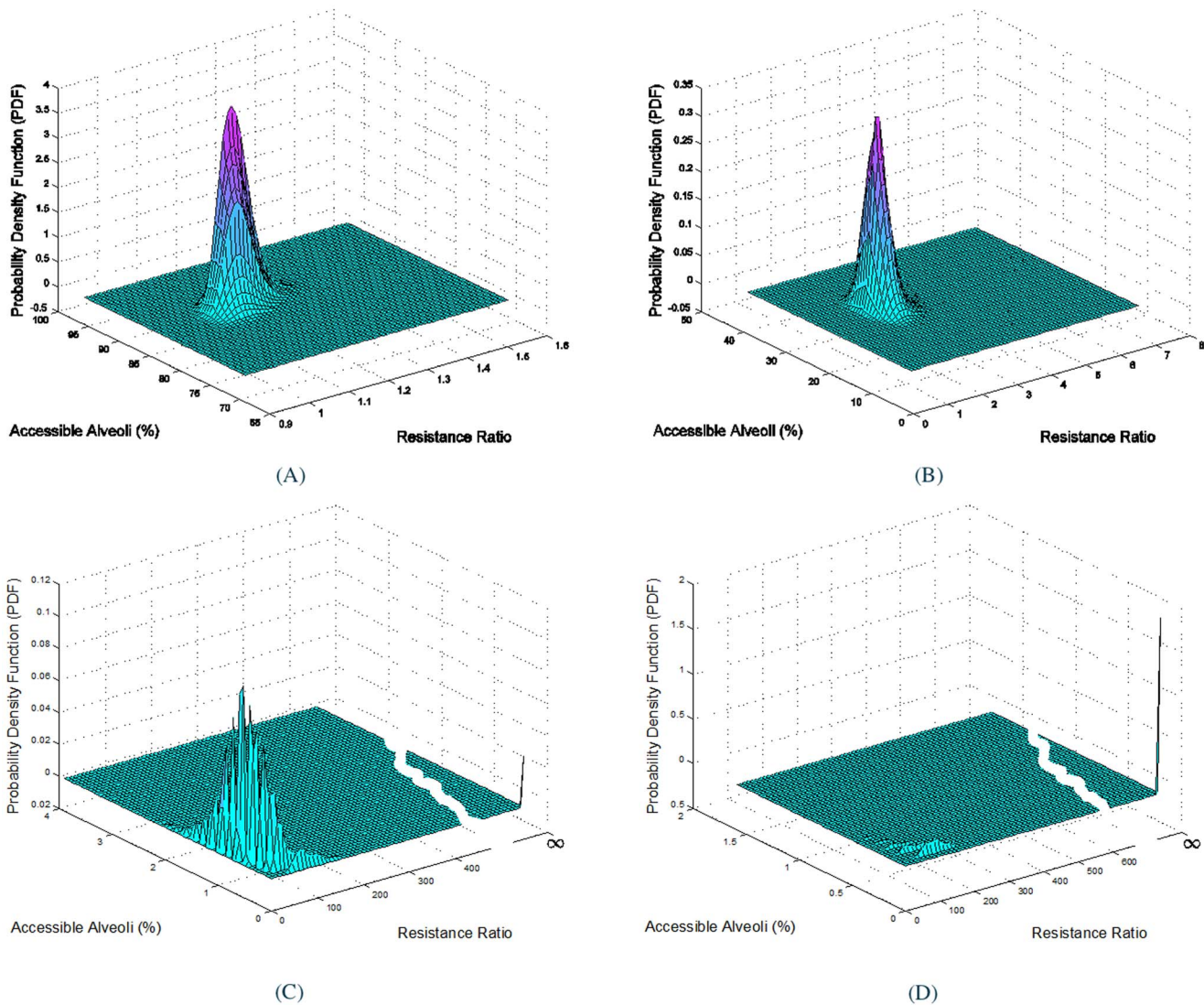
## Results

### Micro-Level Information on Obstructed Bronchioles

In order to spatially locate each bronchiole in a human airway tree we used Kitaoka et al.'s three-dimensional model [18]. Once we receive the imaging data for each voxel, we need to map the values to our model's voxel using spatial location coordinates. Next we use the probability density function that we found in the previous section to determine the corresponding voxel parameters, AA% and resistance ratio, sampled according to the PDFE-2D distributions. The model randomly select a combination of percent accessible alveoli and resistance for each voxel in a way that the total resistance and number of accessible alveoli from these voxels provide the same value as the patient's FEV<sub>1</sub> and FVC. There is a complex calculation taking place in parallel to obtain the rate of flow (FEV<sub>1</sub>) from the total resistance of all the voxels. To find the distribution of voxel parameters we use the Metropolis algorithm to focus the Markov chain to sample many configurations meeting our constraints. This is achieved by choosing the configuration energy provided in Eq(1).

At this stage we have mapped the voxels from imaging data into our model in a way that the total rate of airflow and the accessible alveoli of the lung model resemble the corresponding patient's FEV<sub>1</sub> and FVC values respectively. Once the current state of a CF lung is implemented we can obtain clinically useful information, such as mucus distribution within each voxel or predicted future states of each voxel and associated FEV<sub>1</sub> and FVC values as described in the next section.

Figure 3 displays the flowchart of this process. As shown in this flowchart, once we have all the voxels' parameters, we can select certain voxels for further analysis. We again apply the Metropolis algorithm on the selected voxel in order to visualize the distribution of mucus in its airway tree structure. For the chosen voxel, we have its mucus volume, its resistance and its AA%. We randomly fill out the voxels' airway tree to reach their



**Figure 1. Probability density function for different percent volume of mucus in a small voxel of the lung. (A) 5%, (B) 30%, (C) 65% and (D) 75%** Each voxel represents a subtree from generation 13 to 23 of the binary tree structure of lung. The x-axis is the corresponding airflow resistance and the y-axis shows the percent accessible alveoli. doi:10.1371/journal.pone.0111245.g001

corresponding mucus volume. At each iteration the new resistance and AA% are collected. To move to the neighbor configuration we move a unit volume of mucus in a bronchiole to keep the total mucus volume of the voxel fixed. The state energy we use for this step is as follows:

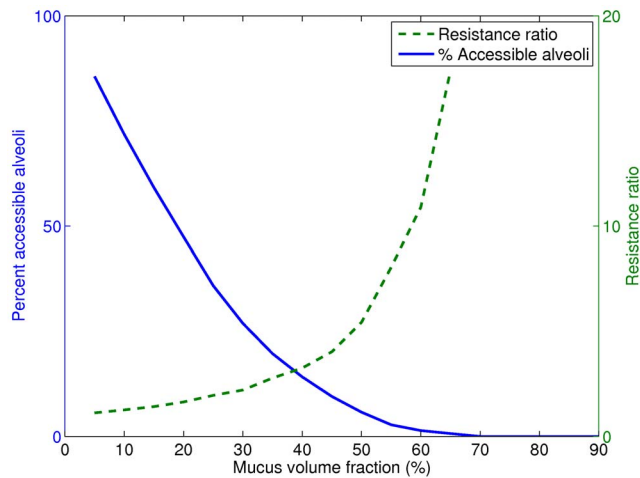
$$E = (\alpha)(\text{Resistance of the current state} - \text{Target voxel's resistance})^2 + (1 - \alpha)(\text{AA\% of the current state} - \text{Target voxel's AA\%})^2 \tag{3}$$

where  $\alpha = 0.5$ . After we obtained enough samples we constructed the corresponding mucus distribution for the selected voxels. Figures 4 displays two examples of the mucus distribution for voxels that contained (5%) and (30%) mucus. As shown in these figures, as the percent mucus increases, the dominantly filled generation moves to bigger bronchioles. The y axis represents % normalized mucus where:

$$\text{Normalized mucus in generation } n = \frac{\left(\frac{\text{Mucus Volume in generation } n}{\text{Total mucus volume in voxel}}\right)}{\left(\frac{\text{Airway volume in generation } n}{\text{Total airway volume in voxel}}\right)} \tag{4}$$

### Predicting future values of FEV<sub>1</sub> and FVC

In this section we use the mucus distribution and growth model presented in [9] to make predictions about the lung functionality of a CF patient. As can be seen in Figure 5 we use the imaging voxels' data and the constant mucus growth rate from [9], or, if available, infection and treatment specific growth rates specific to each voxel to predict the mucus growth in each pocket of infection. The model again resorts to the Monte Carlo method to randomly select the AA% and resistance from our model described in the



**Figure 2. Maximum likelihood of percent of accessible alveoli AA% and airflow resistance for a given mucus volume fraction inside a voxel.** The x-axis is mucus volume% and the two y-axes show the most probable combination of the airflow resistance ratio and percent of accessible alveoli. The resistance value increases as the mucus amount rises. Number of accessible alveoli decreases as the mucus volume grows, which leads to a lower lung functionality test value or FVC value.

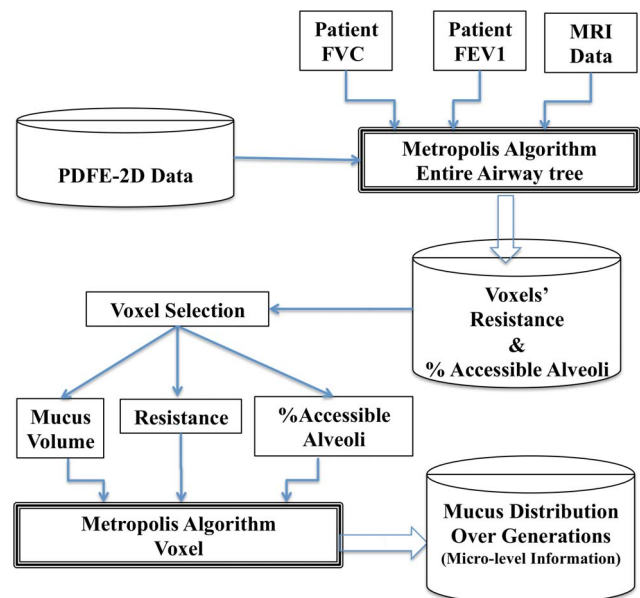
doi:10.1371/journal.pone.0111245.g002

PDFE-2D section and calculates the new FEV<sub>1</sub> and FVC at each iteration and stores their values. After collecting enough samples we can predict the FEV<sub>1</sub> and FVC distributions for a patient after the indicated time period. For example, Figure 6 displays the probability density function of the predicted FEV<sub>1</sub> and FVC. As shown in the figure it is predicted that FEV<sub>1</sub> and FVC of our CF patient are approximately 61% and 78% respectively.

This method can even be applied to follow a patient's progress with out fewer imaging tests. Since most CF patients take spirometry test more often than any imaging tests, we can use our model without recourse to imaging data. This time the model takes only the FEV<sub>1</sub> and the FVC values of a patient as input. After applying the Metropolis algorithm (using additional rearrangements of mucus between voxels) the model can still provide an estimate of the distribution of mucus throughout the airway tree. This can be propagated using the constant any hypothetical growth rates to predict the next spirometric test. We have created a Matlab GUI version of the algorithm; its sample output is shown in Figures 7, 8 and 9.

Since the mean mucus growth rate was estimated based on years of data from many CF patients, such predictions should work reasonably well for a patient between exacerbations. Our model can also be used to predict the progression during exacerbations, albeit with voxel specific growth rates informed by much more data than we currently possess. Using one mean mucus growth rate is a shortcoming not of our algorithm but rather of the paucity of data to which our predictions have been applied. There is every indication that soon we will have reasonable metagenomic [19–22] and metabolomic [23] tools to assess the microbial composition present in a CF lung and will be able to infer growth rates that are specific to the community composition as well as the antibiotic administered. Entering voxel specific growth rates based on more information than we at present possess and using our program to test predictions can move our understanding of the patient's state to a new and quantitative level.

We can improve the accuracy level of the model by extending the resolution of our PDFE-2D distributions. The current PDFEs



**Figure 3. Generating the Micro-level information on obstructed bronchioles.** This flowchart shows how the algorithm produces the mucus distribution of a selected voxel. Using the lung functionality test values: FEV<sub>1</sub> and FVC and the data obtained from MRI lung imaging we can obtain the corresponding mucus distributions among the different generations of lung.

doi:10.1371/journal.pone.0111245.g003

were constructed with 5% bins as described in PDFE-2D model section.

## Discussion

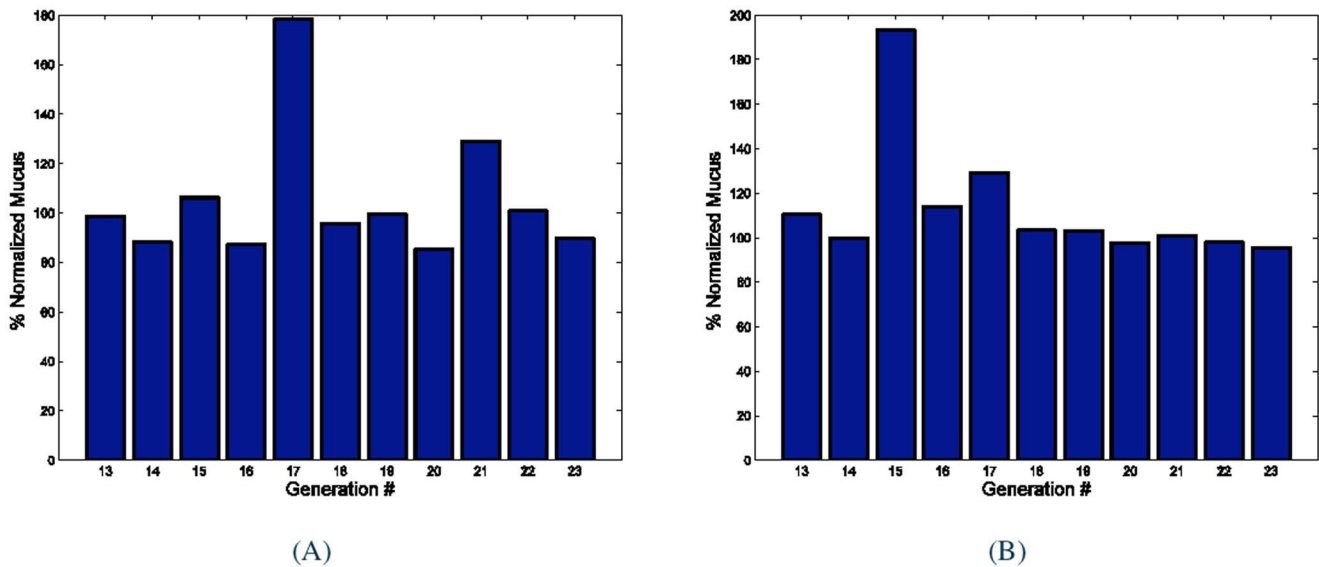
With current research studies about cystic fibrosis, CF treatment is poised for great strides. Non-genetic treatments such as Ivacaftor (trade name Kalydeco, developed as VX-770) only works on patients with a certain mutation of cystic fibrosis which accounts for 4–5% of cystic fibrosis cases [24].

While an eventual cure for the disease by replacement of the defective gene is likely, we expect such treatment not to be available anytime soon. Rather we anticipate that new observational tools such as metagenomics, transcriptomics, metabolomics and MRI imaging coupled with our modeling approach will give the clinician unprecedented ability to follow and treat the disease. The models will provide quantitative predictions of responses to various drug regimens and prescribe adaptively implemented optimal controls for treatment. Predicting the impact of mucus growth on lung functionality will correlate the current stage of the disease with how infection has been propagating throughout different generations of lung. This will enable the physician with a tool to track this propagation of infection.

The various sub-models required here will soon be informed by data characterizing the microbial communities present. Such data comes from metagenomic and transcriptomics analyses of sputum samples and metabolomic analyses of exhaled air. In our current model we predict the dynamic distribution of mucus in a CF lung in the absence of treatment as a stepping stone for eventually eventual treatment and microbial community specific modeling of the treatment response.

According to a research study done by Willner et al. [25] microbial diversity in a CF lung is much higher than suggested by culturing alone. They were able to characterize the diversity of



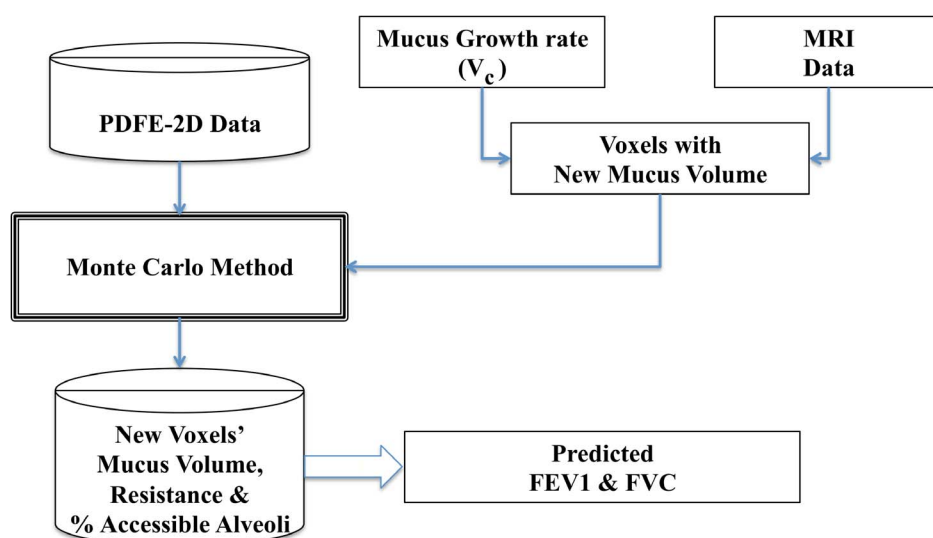


**Figure 4. Mucus distribution among different generations of lung for two selected voxels.** Figure (A) represents a voxel with 5% mucus volume that has 78% accessible alveoli: AA and resistance ratio of 1.22. Figure (B) represents a voxel with 30% mucus that has 19.45% AA and resistance ratio of 4.2. The x-axis is the generation number of the corresponding bronchioles in the selected voxel. The y-axis shows the normalized mucus present, which is equal to the ratio of the mucus amount in a generation to the total mucus volume in a voxel, divided by the ratio of the air volume in that generation to the total voxel volume. doi:10.1371/journal.pone.0111245.g004

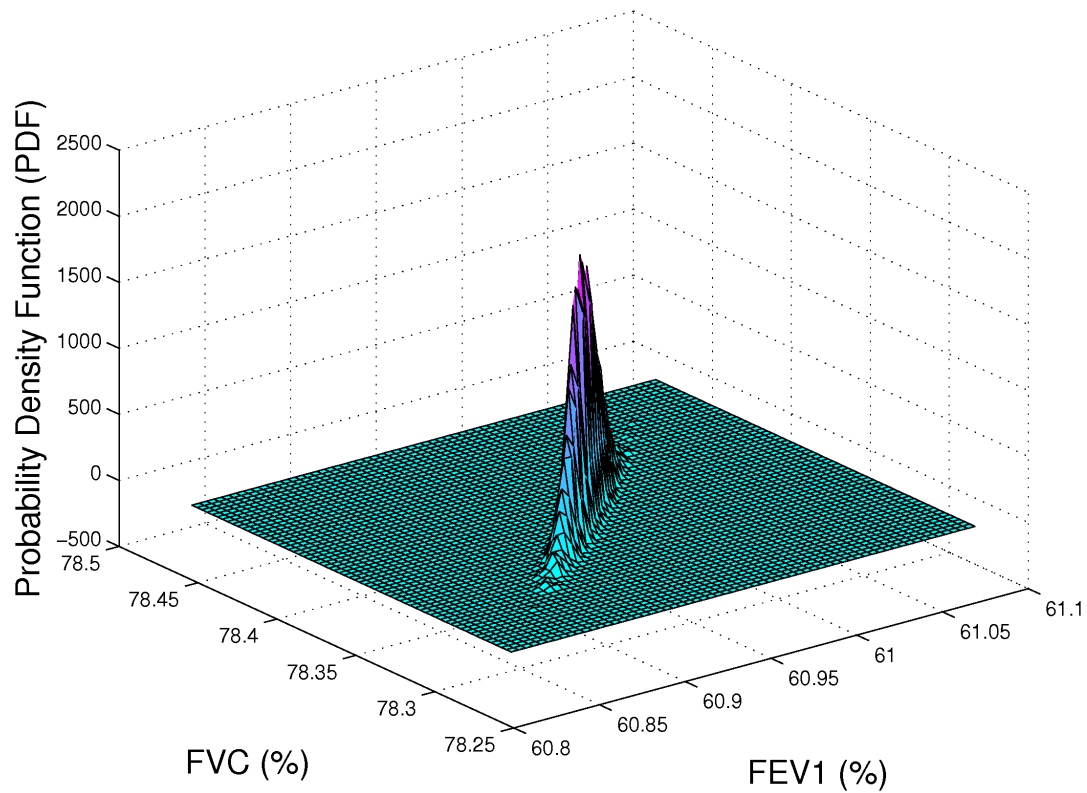
microbial communities in tissue sections from anatomically distinct regions of the CF lung. Their result indicated that microbial communities in the Cystic Fibrosis lung are spatially heterogeneous. This spatial heterogeneity will cause regional differences in microbial biomass and antibiotic resistance. The next version of the model, can use their results to adjust the parameters according to the microbial communities found and the treatment administered (e.g., timing of antibiotic administration, types of antibiotics, steroids, etc). The present model should be taken as a proof-of-concept step toward that goal. This will provide an opportunity for

the researcher, and eventually the clinician, to access a framework for accurate quantitative predictions.

Our methods are completely scalable – the 1.0 cm<sup>3</sup> for the size of one voxel was for illustration. Any resolution scale however forces estimation of the distribution on spatial scales below this resolution. Since estimation of the distribution on finer spatial scales would perforce need the solution of an inverse problem with many possible solutions and since the movement of mucus hinted that the “right answer” would in any case not be a unique distribution, we were led naturally to using Monte Carlo



**Figure 5. FEV<sub>1</sub> and FVC values predictions predicted from the lung model.** This flowchart illustrates the process of how our lung model predicts the FVC and FEV<sub>1</sub> of a CF patient given the previous mucus content of each voxel along with a mucus growth rate. doi:10.1371/journal.pone.0111245.g005



**Figure 6. FEV<sub>1</sub> and FVC probability density function estimation.** The x-axis shows the predicted FEV<sub>1</sub> values and the y-axis is the predicted FVC values for a hypothetical the synthetic lung example described in the text.  
doi:10.1371/journal.pone.0111245.g006

**Personal Data**

Age:

Height (cm):

Male  Female

**Spirometry Data**

FEV1:

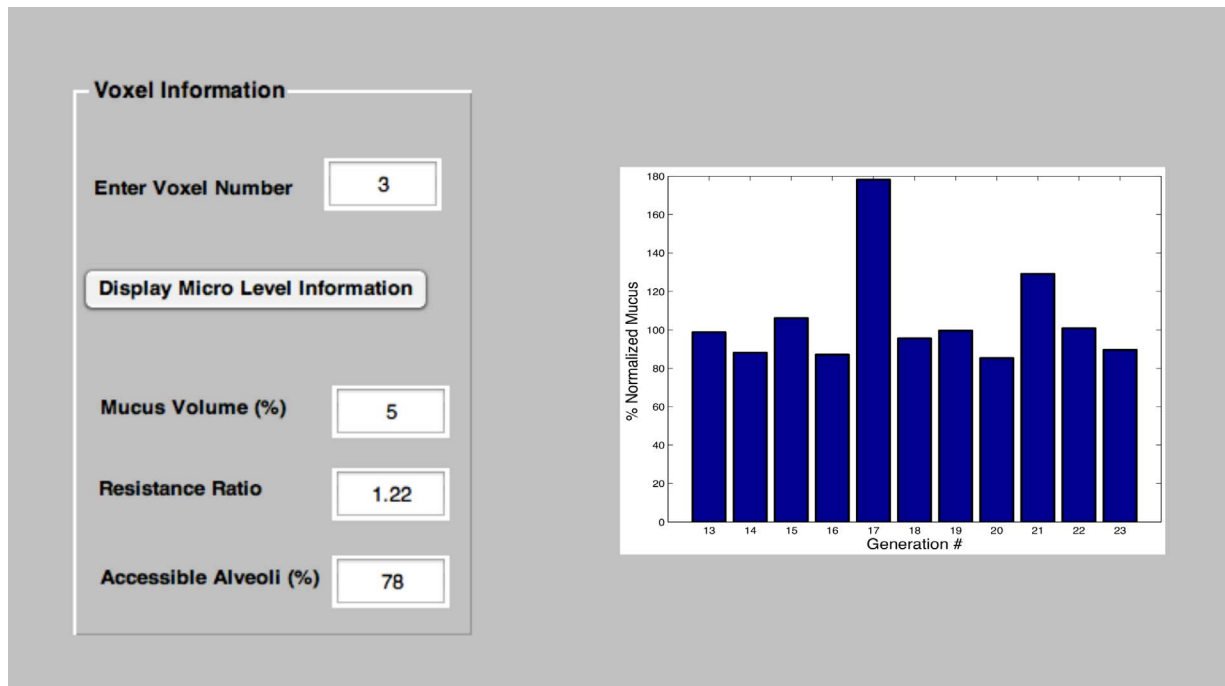
FVC:

**Normalized Spirometry**

FEV1 (%):

FVC (%):

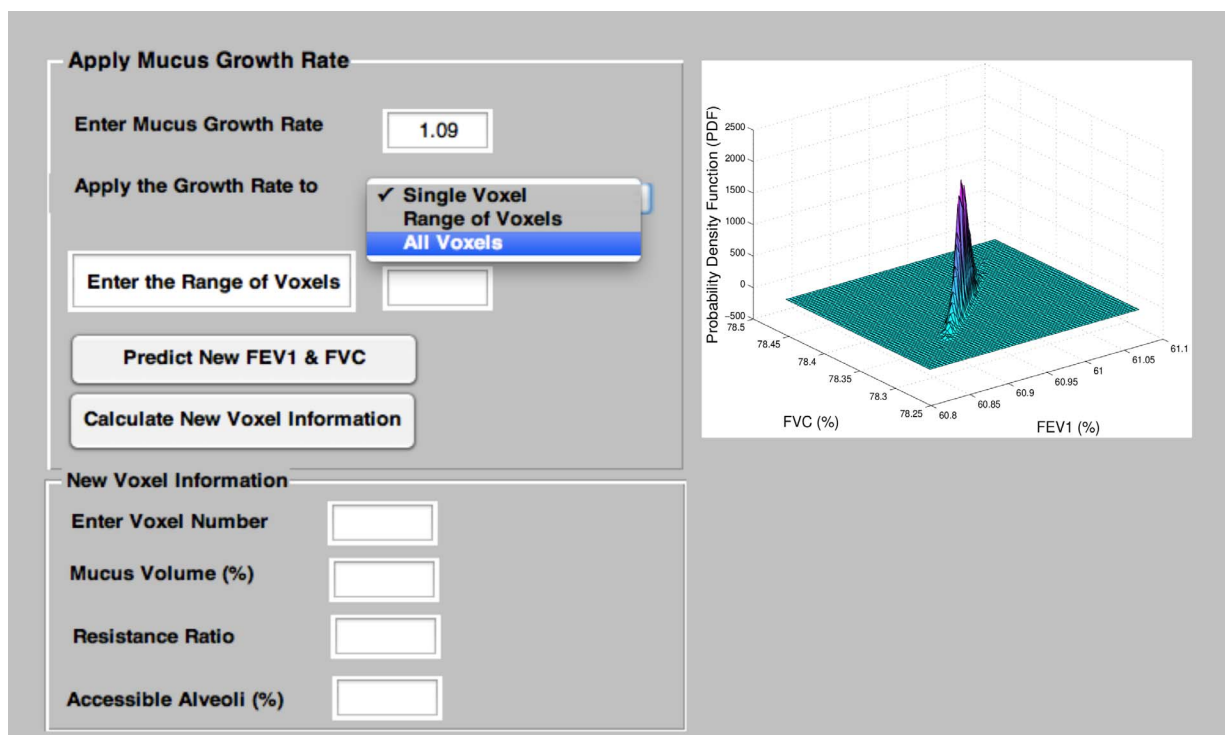
**Figure 7. In this window user can enter the input parameters: patient's age, height, sex and Imaging data.** The Model displays a 3D-lung that contains all the MRI/CT scan voxels.  
doi:10.1371/journal.pone.0111245.g007



**Figure 8. This window shows the mucus distribution among generations.** The user selects a voxel and the model displays its corresponding parameters along with the micro-level information.  
doi:10.1371/journal.pone.0111245.g008

estimation methods for the most likely distribution and the corresponding spirometric observables. The resulting models and algorithms form a clinically useful tool with which to reassess the

various simplest possible sub-models and assumptions used in our work so far. These models can play a crucial role in future treatments of the disease.



**Figure 9. This window provides the user with a probability density function of the predicted FEV<sub>1</sub> and FVC.** User can enter a specific mucus growth rate for different voxels and the new FEV<sub>1</sub> and FVC will be calculated.  
doi:10.1371/journal.pone.0111245.g009



## Supporting Information

**File S1 The supporting file Pseudo.pdf contains the pseudo code of the algorithms and models mentioned in this manuscript.**

(PDF)

## Acknowledgments

This material is based upon work supported by the National Institutes of Health under Grant no. 56586B.

## References

1. Tirouvanziam R (2006) Neutrophilic inflammation as a major determinant in the progression of cystic fibrosis. *Drug News Perspect* 19: 609–614.
2. Koehler DR, Downey GP, Sweezey NB, Tanswell AK, Hu J (2004) Lung inflammation as a therapeutic target in cystic fibrosis. *Amer J of Resp Cell and Mol Biol* 31: 377–381.
3. Eichinger M, Puderbach M, Heussel CP, Kauczor HU (2006) Mri in mucoviscidosis (cystic fibrosis). *Radiologe, German* 46 (4): 275–276.
4. Hopkins SR, Levin DL, Emami K, Kadlecsek S, Yu J, et al. (2007) Advances in magnetic resonance imaging of lung physiology. *J of App Physiol* 102 (3): 1244–1254.
5. Fiel SB, Friedman AC, Caroline DF, Radecki PD, Faerber E, et al. (1987) Magnetic resonance imaging in young adults with cystic fibrosis. *Chest* 91: 181–184.
6. Theilmann RJ, Arai TJ, Samiec A, Dubowitz DJ, Hopkins SR, et al. (2009) Quantitative mri measurement of lung density must account for the change in  $t(2)$  (\*) with lung inflation. *J of MRI* 30: 527–534.
7. Voit EO (2014) Mesoscopic modeling as a starting point for computational analyses of cystic fibrosis as a systemic disease. *Biochim et Biophys Acta* 1844 (1): 258–270.
8. Zarei S, Mirtar A, Andresen B, Salamon P (2013) Modeling the airflow in a lung with cystic fibrosis. *J of Non-Equil Thermodyn* 38 (2): 119–140.
9. Zarei S, Mirtar A, Rohwer F, Conrad DJ, Theilmann RJ, et al. (2012) Mucus distribution model in a lung with cystic fibrosis. *J of Comp and Math Meth in Med*.
10. Weibel ER (1991) Design of airways and blood vessels considered as branching trees. *Lung: Sci Found* 1: 1061–1071.
11. Altmeier WA, McKinney S, Glenny RW (2005) A fractal analysis of the radial distribution of bronchial capillaries around large airways. *J of App Physiol* 98: 850–855.
12. Glenny RW, Bernard SL, Robertson HT (2000) Pulmonary blood flow remains fractal down to the level of gas exchange. *J of App Physiol* 89: 742–748.
13. West BJ (1990) Physiology in fractal dimensions: error tolerance. *Anna of Biomed Engi* 18 (2): 135–149.
14. Richey M (2010) The evolution of markov chain monte carlo methods. *Amer Math Mon* 117 (5): 383–413.
15. Saeta PN (2004). The metropolis algorithm: Statistical systems and simulated annealing. URL <http://saeta.physics.hmc.edu/courses/p170/Metropolis.pdf>.
16. Salamon P, Sibani P, Frost R (2002) Facts, Conjecture, and Improvements for Simulated Annealing. Philadelphia: Society for Industrial and Applied Mathematics.
17. Marieb E, Hoehn K (2010) Human Anatomy and Physiology. Prentice Hall PTR. 8th ED.
18. Kitaoka H, Suk B (1997) Branching design of the bronchial tree based on a diameter-flow relationship. *J of App Physiol* 82: 968–976.
19. Willner D, Furlan M, Haynes M, Schmieder R, Angly FE, et al. (2009) Metagenomic analysis of respiratory tract dna viral communities in cystic fibrosis and non-cystic fibrosis individuals. *PLoS One* 4 (10): e7370.
20. Lim YW, Evangelista JS, Schmieder R, Bailey B, Haynes M, et al. (2014) Clinical insights from metagenomic analysis of sputum samples from patients with cystic fibrosis. *J of Clin Microbi* 52 (2): 425–437.
21. Lim YW, Schmieder R, Haynes M, Furlan M, Matthews TD, et al. (2013) Mechanistic model of rothia mucilaginosa adaptation toward persistence in the cf lung, based on a genome reconstructed from metagenomic data. *PLoS one* 8 (5): e64285.
22. Lim YW, Schmieder R, Haynes M, Willner D, Furlan M, et al. (2013) Metagenomics and metatranscriptomics: Windows on cf-associated viral and microbial communities. *J of Cystic Fibrosis* 12 (2): 154–164.
23. Whiteson KL, Meinardi S, Lim YW, Schmieder R, Maughan H, et al. (2014) Breath gas metabolites and bacterial metagenomes from cystic fibrosis airways indicate active ph neutral 2, 3-butanedione fermentation. *ISME J*.
24. McPhail G, Clancy J (2013) Ivacaftor: The first therapy acting on the primary cause of cystic fibrosis. *Drugs of today* 49: 253–260.
25. Willner D, Furlan M, Haynes M, Schmieder R, Lim YW, et al. (2011) Spatial distribution of microbial communities in the cystic fibrosis lung. *ISME J* 6: 471–747.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author Contributions

Conceived and designed the experiments: SZ AM FR PS. Performed the experiments: SZ AM FR PS. Analyzed the data: SZ AM FR PS. Contributed reagents/materials/analysis tools: SZ AM FR PS. Wrote the paper: SZ AM FR PS.