Cardiopulmonary responses to maximal aerobic exercise in patients with cystic fibrosis

Craig A. Williams[ID]¹*, Kyle C. A. Wedgwood²,³, Hossein Mohammadi²,⁴, Katie Prouse², Owen W. Tomlinson[ID]¹, Krasimira Tsaneva-Atanasova²,³,⁴

1 Children's Health and Exercise Research Centre, Sport and Health Sciences, University of Exeter, Exeter, United Kingdom, 2 Department of Mathematics and Living Systems Institute, University of Exeter, Exeter, United Kingdom, 3 Centre for Biomedical Modelling and Analysis, University of Exeter, Exeter, United Kingdom, 4 EPSRC Centre for Predictive Modelling in Healthcare, University of Exeter, Exeter, United Kingdom

* c.a.williams@exeter.ac.uk

Abstract

Cystic fibrosis (CF) is a debilitating chronic condition, which requires complex and expensive disease management. Exercise has now been recognised as a critical factor in improving health and quality of life in patients with CF. Hence, cardiopulmonary exercise testing (CPET) is used to determine aerobic fitness of young patients as part of the clinical management of CF. However, at present there is a lack of conclusive evidence for one limiting system of aerobic fitness for CF patients at individual patient level. Here, we perform detailed data analysis that allows us to identify important systems-level factors that affect aerobic fitness. We use patients' data and principal component analysis to confirm the dependence of CPET performance on variables associated with ventilation and metabolic rates of oxygen consumption. We find that the time at which participants cross the gas exchange threshold (GET) is well correlated with their overall performance. Furthermore, we propose a predictive modelling framework that captures the relationship between ventilatory dynamics, lung capacity and function and performance in CPET within a group of children and adolescents with CF. Specifically, we show that using Gaussian processes (GP) we can predict GET at the individual patient level with reasonable accuracy given the small sample size of the available group of patients. We conclude by presenting an example and future perspectives for improving and extending the proposed framework. The modelling and analysis have the potential to pave the way to designing personalised exercise programmes that are tailored to specific individual needs relative to patient's treatment therapies.

Introduction

Cystic fibrosis (CF) is the most common life shortening genetic disease in the Caucasian population, affecting nearly 11,000 individuals in the United Kingdom (UK) [1] and predominantly manifests itself throughout the respiratory, digestive and reproductive systems of the human body. The genetic mutation responsible for CF results in reduced trans-epithelial chloride transport, and increased sodium and water absorption, thus reducing the hydration status of

the mucosal lining of the airway and digestive systems. The resultant viscous mucus is liable to increased infection and further inflammation and a progressive decline in lung function [2]. Currently there is no cure for CF and therefore CF is a diseased that is managed with appropriate medication, nutrition, physiotherapy and exercise. Enhanced aerobic fitness as represented by maximal oxygen consumption ($\dot{V}O_{2max}$) is associated with lower risk of hospitalisation, increased exercise tolerance, reduced residual volume, increased endurance of the respiratory muscles, enhanced sputum expectoration and decreased rate of decline in pulmonary function [3–7]. Furthermore, individuals with CF possessing a higher $\dot{V}O_{2max}$ are shown to have a reduced mortality risk, as Nixon et al. [8] reported that adults with a $\dot{V}O_{2max}$ greater than 82% of their predicted value had an 83% 8-year survival rate, compared to just 28% 8-year survival rate for patients with a $\dot{V}O_{2max}$ less than 58% of their predicted value [8]. This enhanced survival has also been reported in a paediatric cohort, with a 100% survival rate after 7 years in children with CF with a $\dot{V}O_{2peak}$ over 45 ml·kg^{-1}·min^{-1} [9].

Given the importance of identifying and monitoring aerobic fitness, cardiopulmonary exercise testing (CPET) is recommended to take place on at least an annual basis [10], to provide a clinically useful prognostic evaluation of a patient's functional capabilities. Even in mild to moderate severity of CF, patients are known to demonstrate impairments in cardiac and respiratory functions leading to exercise intolerance which are responsive to exercise training, and can result in an enhanced quality of life, increased physical function and increased life expectancy [11, 12].

Effective management of the disease is of critical importance due to an aging CF population group (median predicted survival of children born with CF in the UK is currently 45 years [1]) and high medical care costs [13]. Exercise is widely acknowledged as a key management strategy for CF, supported by mechanistic data on the systemic effects of exercise at the cellular level *in vivo* in young patients with CF [6–8, 11, 14]. However, an integrated systems level understanding of the limitations of aerobic fitness for CF patients is lacking. Measurement techniques that do exist to quantify within-organ, real-time perfusion and intracellular oxygenation are invasive and unethical for use with paediatric patients, and current animal model research provides limited direct relevance to paediatric pathology. In clinical practice, there is significant interaction between cardiac, pulmonary and musculoskeletal function, which can result in the functional improvement in one part of the combined system, but detrimental effects on others [11]. Clinicians therefore inevitably have to adopt very imprecise guidelines related to exercise prescription [15].

The use of modelling and simulation tools in clinical medicine is currently the subject of intense research interest both in the UK and internationally [16–20], and the adoption of a systems biomedicine approach to build and validate novel multi-scale, organ-level, integrated, re-usable and re-deployable models represents a paradigm shift in biomedical modelling and simulation. There are numerous organ level models in existence [21–24], however, to date there have been limited attempts to either integrate these or to apply them to real clinical applications. There is ongoing basic science and clinical trial work providing data on the micro [25] and macrovascular [26] changes associated with exercise. These data, although important, have yet to be integrated quantitatively with other data streams. In particular, there has been no previous work on the use of predictive modelling and simulation technologies for developing treatment strategies for CF patients.

Therefore the aim of this study is twofold. Firstly, to analyse the physiological responses to progressive exercise in patients with CF, with a view of determining predictors of performance. Secondly, to develop a surrogate (statistical) model that allows the evaluation of how CF impairs exercise tolerance relative to increasing ventilatory and metabolic demands.

Specifically, how various ventilatory parameters quantitatively affect the gas exchange threshold (GET) for a group of patients is explored.

Materials and methods

This study was a retrospective analysis of existing CPET data from 15 children and adolescents with CF. Original ethics approvals was approved by South West NHS Research Ethics Committees [10/H0107/78; 13/SW/0166; 14/SW/0061] and written informed consent and assent was obtained from parent(s)/guardian(s) and participants, respectively in accordance with the Declaration of Helsinki.

CPET data

As part of original studies, all participants performed a valid [27] and reliable [28] combined ramp incremental and supramaximal (S_{max}) CPET to determine $\dot{V}O_{2max}$ and the GET. This protocol was performed on an electronically braked cycle ergometer, and required patients to perform an initial exhaustive ramp incremental test at a pre-determined rate between 10–25 $W \cdot min^{-1}$, in order to elicit exhaustion in approximately ten minutes [29]. After a 3-min warm-up at 10–20 W, participants completed this incremental test to the point of volitional exhaustion, maintaining a cadence of 70–80 rpm throughout. Exhaustion was defined as a 10 rpm drop in cadence for five consecutive seconds, despite strong verbal encouragement. Active (5-min cycling at 20 W) and then passive seated recovery (10 min) then preceded the S_{max} bout. S_{max} verification consisted of a 3-min warm-up (10–20 W), followed by a "step" transition to a constant work rate corresponding to 110% peak power output [30] obtained during the ramp incremental phase.. Upon volitional exhaustion (defined previously), a 5-min active recovery (slow cycling at 20 W) concluded the combined CPET session. $\dot{V}O_{2max}$ and the GET were subsequently identified for all participants using established methods which are validated for use in CF [31].

Models and simulations

Simulations are widely used in various fields of science and engineering because conducting physical experiments is too costly, or highly time-consuming, or even impossible in some cases [32]. In the case of CPET in CF patients, there are also ethical considerations, since the test adds to the treatment burden many children and adolescents with CF already face.

Often, a primary goal of using model simulations is to perform quantitative studies such as uncertainty quantification or sensitivity analysis. Such studies are crucially important in biomedicine, since there exists significant variation both between and within patient groups. Through understanding and quantification of the uncertainty within the mathematical models, outcomes of patient-specific interventions can be better predicted. However, such investigations require a large number of runs that makes it impractical if each run takes more than a few seconds. To cope with this difficulty, one can use *emulators*, also known as *surrogates*, or *metamodels* or *response surfaces* [33]. These provide a fast approximation of the input/output relation governed by the underlying simulator. The most important classes of surrogate models have been described elsewhere [34–36].

The surrogate model employed in this study is based on Gaussian processes (GPs), which have become increasingly popular over the last two decades [33]. A Gaussian Process (GP) defines a probability distribution over functions where the true function is considered as a particular sample path. GPs have been used in a wide range of applications from wireless communication, to obtain position estimates for a mobile user [37]; metallurgy, to model the

development of microstructure [38]; and in biology, to describe gene regulatory processes and cell growth [39–41]. A GP defines a probability distribution over functions which is fully specified by its mean, μ, and covariance K [42], which are both functions of the input variables: μ = μ(**x**), $K = K($**x,x'**$)$. While μ(**x**), could be any function (though in practice is often chosen to have a polynomial dependence on **x)**, the covariance function is required to be positive semi-definite [42]. Specifically, if Y is a GP defined on the space of input variables denoted by D, then we write:

$$Y \sim GP(\mu, K) : \ \mu(\mathbf{x}) \ = \ \mathrm{E}[Y(\mathbf{x})], \ K(\mathbf{x}, \mathbf{x}') \ = \ \mathrm{Cov}(Y(\mathbf{x}), Y(\mathbf{x}')), \ \forall \ \mathbf{x}, \mathbf{x}' \in D \subset \mathbb{R}^d$$

The above can be regarded as "prior" distribution over function spaces. This can be seen more clearly in Fig 1(a). In this subfigure, which shows a generic example of a GP, the bold red line is the 'true' function f. Note that the true function is unknown—our aim is to construct a model that approximates it. The thin grey lines are sample functions of a GP distribution with μ(**x**) = 0 and a squared-exponential covariance function given by:

$$K(\mathbf{x}, \mathbf{x}') \ = \ \sigma^2 \prod_{i=1}^{d} exp\left(-\frac{|x_i - x_i'|^2}{2\Phi_i^2} \right),$$

Which is a popular choice in GP modelling. Here, σ controls vertical variability of *sample functions* and $\Phi_i > 0$, $i = 1, \ldots, d$, governs the degree of smoothness of them along the input dimension i. These parameters are usually unknown and estimated via maximum likelihood method [42]. For exposition purposes, the example plot is restricted to the case $d = 1$, but the approach is unchanged for $d > 1$.

Thus far, we have defined the prior distribution for the GP. It is clear from Fig 1(a) that, in general, prior distributions are unlikely to provide a good approximation to the true function f. We can compute the GP "posterior" distribution by incorporating data points (known as "training data") obtained from evaluating f at specific points, following a Bayesian framework. The resulting posterior distribution of the GP, conditioned on the data, will be much closer to the true function.
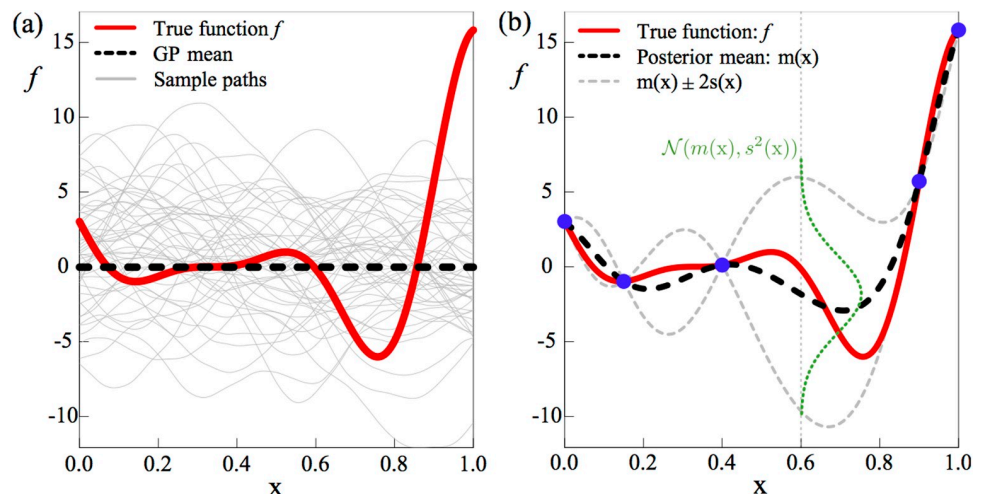
Fig 1. GP is a probability distribution over functions. The thick solid red line is the true function f and the thick dashed black line is the GP "prior". (a) Thin grey lines show sample functions of the GP. (b) Blue bullets indicate five data points sampled from f. The GP distribution is updated using these sample data points. The thin grey dashed lines show $m(\mathrm{x}) \pm 2s(\mathrm{x})$.

Let $\mathbf{y} = \{f(\mathbf{x}^1), \ldots, f(\mathbf{x}^n)\}$ be a set of function evaluations (or observations) at n locations $\mathbf{X} = \{\mathbf{x}^1, \ldots, \mathbf{x}^n\}$. Here, function evaluations correspond to the GET location for a set of patients during the CPET. Predicting with GP is obtained by conditioning it Y on sample points the observations, i.e. $\Omega = \{X, y\}$ $Y(x)|Y(x^1) = f(x^1), \ldots, Y(x^n) = f(x^n)$. For any (new) $z \in D$, this conditional distribution has a normal distribution with the mean m(z) and variance $s^2(z)$ which are used as the prediction and the associated uncertainty at z posterior, respectively. They are expressed by distribution of $Y(z)|\Omega$ has a normal distribution with the following mean, m, and variance, s^2:

$$m(\mathbf{z}) = \mu(\mathbf{z}) + K(\mathbf{z}, \mathbf{X})^{\mathrm{T}} K(\mathbf{X}, \mathbf{X})^{-1}(\mathbf{y} - \boldsymbol{\mu}) \tag{1}$$

$$s^2(\mathbf{z}) = K(\mathbf{z}, \mathbf{z}) - K(\mathbf{z}, \mathbf{X})^{\mathrm{T}} K(\mathbf{X}, \mathbf{X})^{-1} K(\mathbf{z}, \mathbf{X}) \tag{2}$$

where $^{\mathrm{T}}$ denotes the transpose operator, $^{-1}$ is the inverse operator and $\boldsymbol{\mu} = \mu(\mathbf{X})$ is the vector of the mean function at $\mathbf{X}$. In addition, $K(\mathbf{z},\mathbf{X})$ and $K(\mathbf{X},\mathbf{X})$ are the covariance vector between Y ($\mathbf{X}$) and $Y(\mathbf{z})$ and the covariance matrix between the observations. Fig 1(b) shows an example of incorporating sample points to update the prior distribution shown in Fig 1(a). In this generic example, the function f is evaluated at five distinct values of $\mathbf{x}$, and the mean and variance of the GP are updated using (1)-(2).

At the evaluated points, indicated in blue (colour online), the true value of f is known and so the variance of the GP at these points vanishes and m($\mathbf{x}$) = $f(\mathbf{x})$. In between these points, the variance increases, dependent on the distance (in terms of $\mathbf{x}$) from a sampled point. The mean of the GP, shown by the thick black dashed line now approximates the true function much more closely (recall that the prior mean function in Fig 1(a) was zero everywhere), and matches exactly at the evaluated points. The approximation can be further improved by incorporating more data points (function evaluations), particularly around those input values for which the prediction variance $s^2(\mathbf{x})$ is high.

The data analysis was performed using Python (Anaconda Software Distribution. Version 2–2.4.0. Continuum Analytics, 2016. URL https://continuum.io) and MATLAB and Statistics Toolbox Release 2016b, The MathWorks, Inc., Natick, Massachusetts, United States. The GP model simulator was implemented in R (R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.)

Results

Data analysis

To facilitate understanding, we first plot in Fig 2(a), raw data displaying the performance of the participants. The work rate for each participant is increased at a rate that is either, a) dependent on their performance in previous tests, or b) when a prior test is unavailable, at a rate that is predicted to elicit exhaustion in approximately ten minutes [29]. This is done in order to keep the expected duration of the test comparable to other participants. Note that this means that the total energy expended by a given participant is not based on the duration of the test alone. In Fig 2(b), we show how participant age affects overall test performance. We observe a correlation between the two: the worst performing participants tend to be the youngest, but this effect is insignificant at older ages. The colour coordination used in this figure (red-worst performance → blue-average performance → green-best performance) will be used throughout the remainder of this section, where performance is quantified by the total energy transferred during the test.
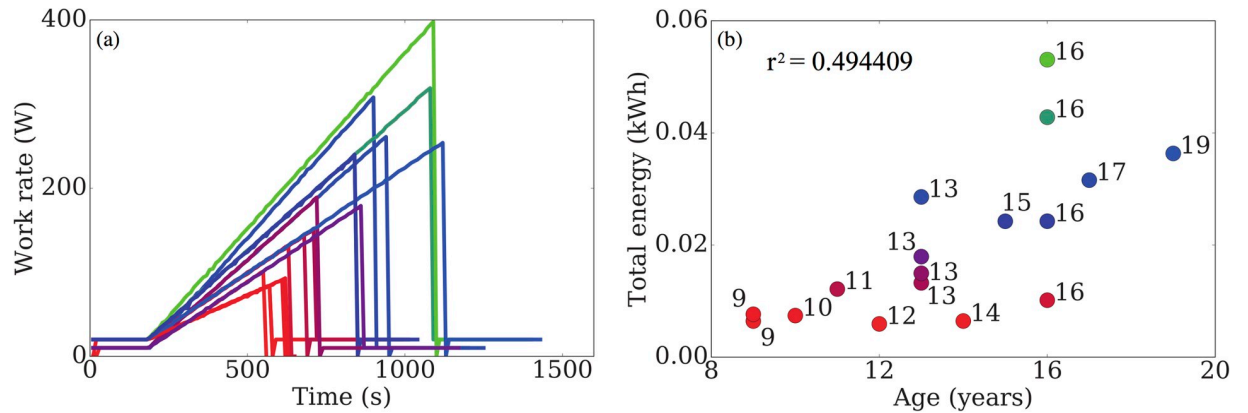
Fig 2. (a) The work rate for each participant is increased at a rate dependent on their past test performance. (b) Participant age is correlated with test performance for young participants, but not for older ones.

In Fig 3, we plot the ratio of $\dot{V}O_2$ over total ventilation ($\dot{V}_E$) with respect to time. The markers on each of the time traces indicate the time of volitional exhaustion for that participant. There are two features that stand out from this figure. Firstly, participants who perform better have higher $\dot{V}O_2/\dot{V}_E$ ratios, suggesting that their oxygen uptake is more efficient than their poorer performing counterparts. Secondly, in the recovery phase of the test (5 minutes following volitional exhaustion) better performing participants exhibit a sharp decrease in $\dot{V}O_2/\dot{V}_E$, which is not observed in the poor performance group. Again, this suggests a more efficient utilisation amongst the former group and that exhalation of CO_2 is perhaps more significant to total breathing following the test.

We next examine the effect that breathing patterns have on participant performance. Two classical prognostic measures used for patients with cystic fibrosis are the forced vital capacity (FVC), and the forced expiratory volume in one second (FEV$_1$). These measures have been shown to be well correlated with mortality and overall fitness of CF patient groups [43–45]. In Fig 4, we demonstrate how these metrics are correlated with performance in the CPET test. In
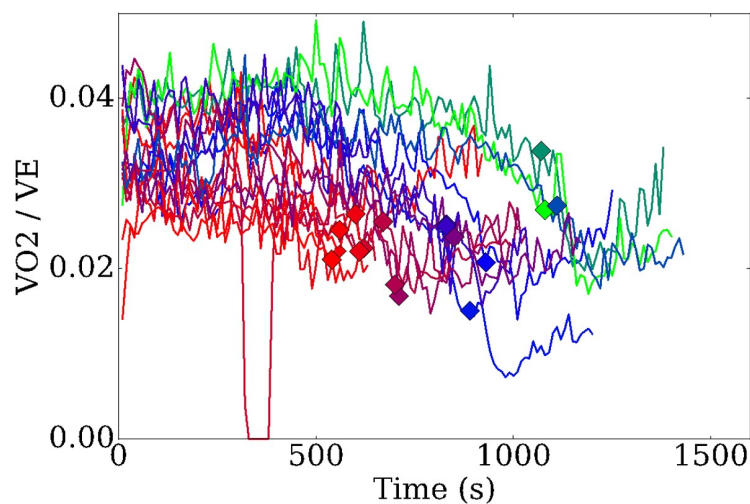
Fig 3. Ratio of oxygen utilisation and total breathing throughout the test. Markers indicate the volitional exhaustion times for each participant.
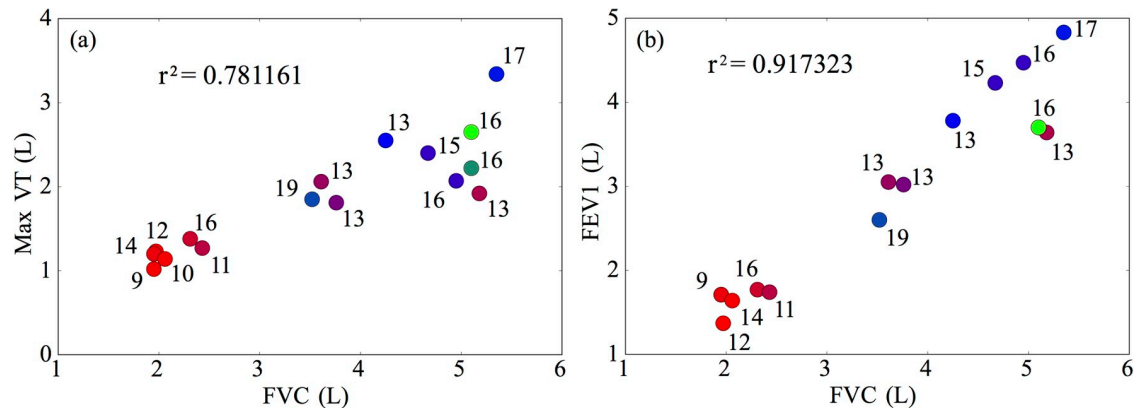
Fig 4. (a) Correlation of FEV_1 with the maximal tidal volume achieved throughout the test. (b) Correlation between FEV_1 and FVC is high. Note that, although FVC and FEV_1 are good predictors of poor test performance, they are unable to distinguish better performing participants.

https://doi.org/10.1371/journal.pone.0211219.g004

Fig 4(a), we observe good correlation between FVC and the maximum tidal volume (VT) of breathing achieved throughout the test. This is unsurprising since participants are likely to be trying to maximise their breathing depth close to their exhaustion point. However, notice that, although the group with low FVC performed poorly, this measure was unable to separate other participants. Fig 4(b) reiterates this result and also highlights the high correlation between FEV_1 and FVC.

In order to better classify the performance of the participants, we must instead look for other factors. In Fig 5, we present the total breathing rate against the oxygen consumption throughout the test. In Fig 5(a), we find a strong relationship between test performance and respiratory pattern. Note that the curvature of the graphs suggests that an exponential fit, rather than a linear one, is most appropriate for these data. In order to test this, we take logarithms of the data and perform a linear regression, ignoring the first 180s of the test since participants are here in the warm up phase (work rate is not increasing) and the final 60s of the data prior to volitional exhaustion, since participants pass their respiratory compensation point, inducing hyperventilation and erratic breathing. The results of the fit are shown in Fig 5(b), where we can clearly see the dependence of performance on breathing pattern.
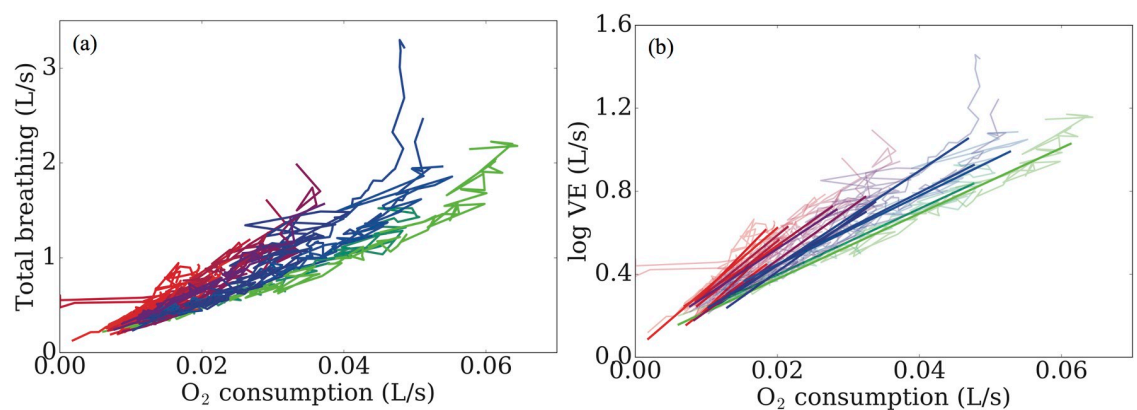
Fig 5. (a) Total ventilation plotted against oxygen utilisation. We observe that breathing pattern is strongly correlated with test performance. (b) Exponential curves are fitted through the raw data, further highlighting this dependence.

https://doi.org/10.1371/journal.pone.0211219.g005
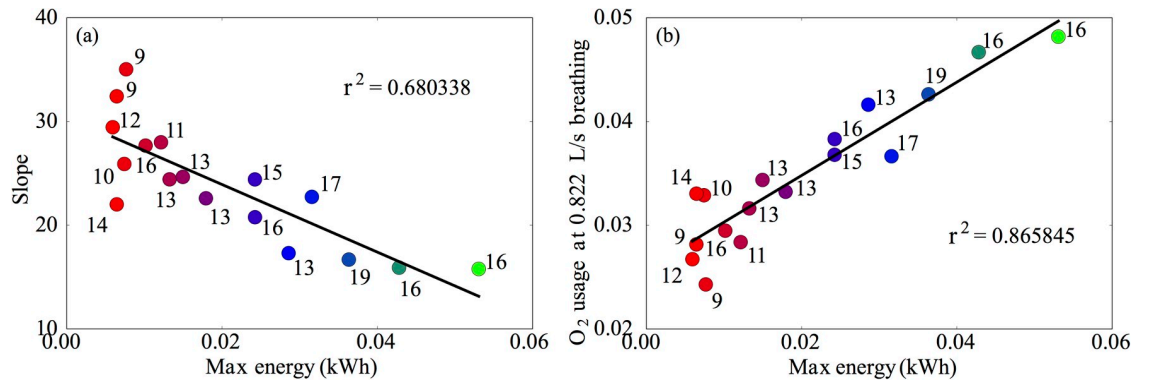
Fig 6. Slope of the fitted curves (log$\dot{V}_E$ against $\dot{V}O_2$) from Fig 4(b) plotted against the total energy transfer during the test. We find a relatively poor characterisation of the variance between performances. (b) By instead plotting the oxygen consumption at a fixed rate of breathing, we better capture differences in performance.

From the fitted curves, we can further explore the dependence of breathing patterns on performance. Firstly, in Fig 6(a), we depict the slope of the fitted curve against the total energy transfer. We find that the slope of the curve of log$\dot{V}_E$ against $\dot{V}O_2$ alone does not capture all of the variation in energy, which is highlighted by the relatively low coefficient of determination (0.68). Instead, we plot in Fig 6(b) the oxygen consumption at a fixed rate of breathing against the total energy. Here, we find a very good characterisation of the overall performance, with a much higher coefficient of determination (0.86), confirming that those who utilise oxygen more efficiently perform better.

Next, we examine the specific patterns of breathing exhibited by the participants, in particular, focussing on breathing depth and frequency. Initial characterisations of these patterns seem to provide little information, as indicated in Fig 7(a). However, when we now also include dependence of oxygen consumption, we find a near perfect classification of participants into the lowest performing groups, the best performing groups and the middle group. These data are shown in Fig 7(b). Note that in this figure, the trajectories appear to be evolving on a planar manifold, suggesting significant co-dependence between these three variables.

Given that there appears to be co-dependence between the variables used in Fig 7(b), a natural next step is to use principal component analysis (PCA) to account for these dependencies. By projecting the data onto their principal components, we show in Fig 8 how well these
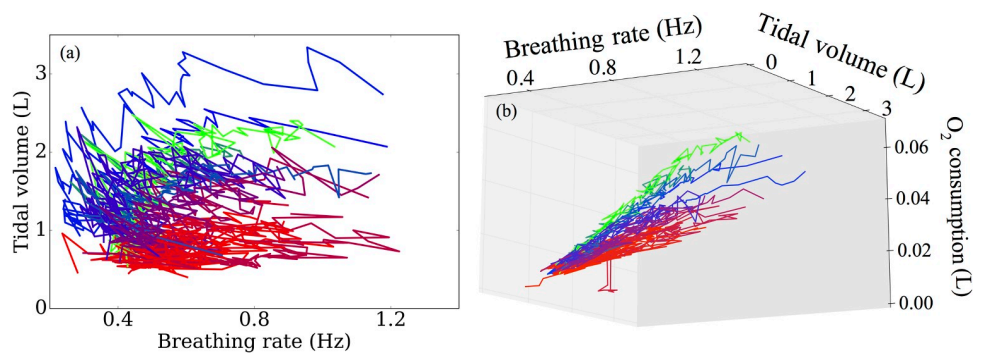
Fig 7. (a) Breathing patterns subdivided into the breathing rate and tidal volume. These data appear uninformative for predicting test performance. (b) With the additional inclusion of the oxygen consumption at a fixed rate of breathing, we find that these variables now almost perfectly capture variation in participant performance.
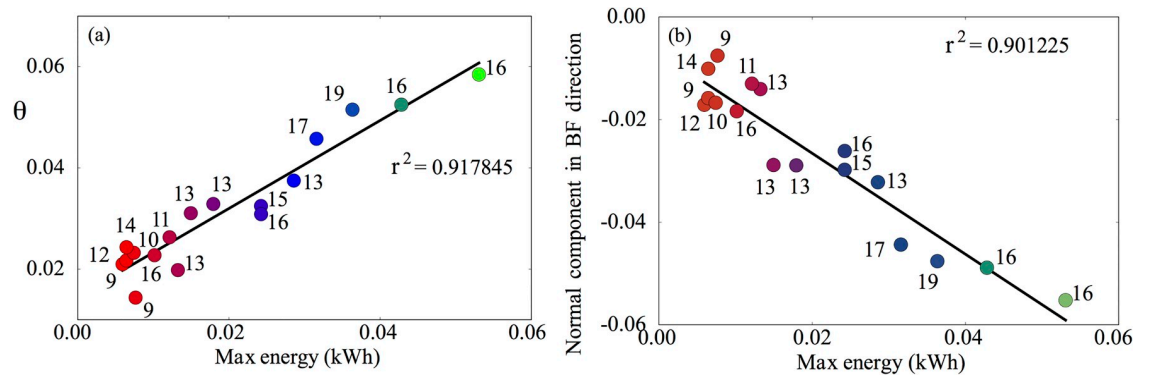
Fig 8. The first principal component obtained via PCA accounts for over 90% of the variation in test performance. (b) Similar levels of variance are accounted for by taking only the normal component of the first principal component, θ, in the breathing frequency direction.

components capture the variation in participant performance. Given that there are only three independent variables in our analysis, it is convenient to use spherical polar coordinates to show how these quantify performance. The first of these components, θ, captures over 90% of the variation in performance (Fig 8a), as does the normal component in the direction of breathing frequency (Fig 8b). These results further indicate the importance of breathing frequency, together with co-variation of oxygen consumption and tidal volume as predictors of test performance.

Gas exchange threshold

Under steady state levels of exercise, the metabolic rate of production of CO_2 is assumed to be proportional to the utilisation rate of O_2 via the cellular respiratory quotient, since (after the initial rest-work transition) adenosine triphosphate (ATP) is replenished primarily via aerobic metabolism pathways. As the work rate increases, this pathway becomes unable to supply sufficient ATP to satisfy the required amount of energy and anaerobic pathways have to contribute to overcome the shortfall. In so doing, they increase the levels of metabolic waste products, such as lactate and also increase the overall production rate in CO_2. The point at which this occurs is known as the anaerobic threshold (AT) or sometimes can be referred to as the lactate threshold. The lactate threshold can be estimated non-invasively by the determination of the GET [46]. Within this study, the GET was successfully identified in all CPETs.

Anaerobic thresholds are well correlated with overall performance, as shown in Fig 9(a), due to the fact that the anaerobic pathways are less efficient at producing ATP and because build-up of lactate contributes significantly to fatigue. One of the major contributing factors in defining GET is $\dot{V}O_{2max}$, since this is indicative of the limit of the rate of oxidative phosphorylation. It thus comes as little surprise that $\dot{V}O_{2max}$ is the best single predictor of CPET performance, as shown in Fig 9(b).

Gaussian processes-based modelling

Having completed a detailed data analysis and identified candidate predictors of CPET performance we next use these predictors to inform a Gaussian Process-based model. In our modelling, we attempt to describe the influence of breathing patterns and $\dot{V}O_2$ on the GET, since this is shown to correlate well with overall test performance (see Fig 9). In mathematical terms, we treat GET as our scalar output variable, with input variables comprising: baseline breathing
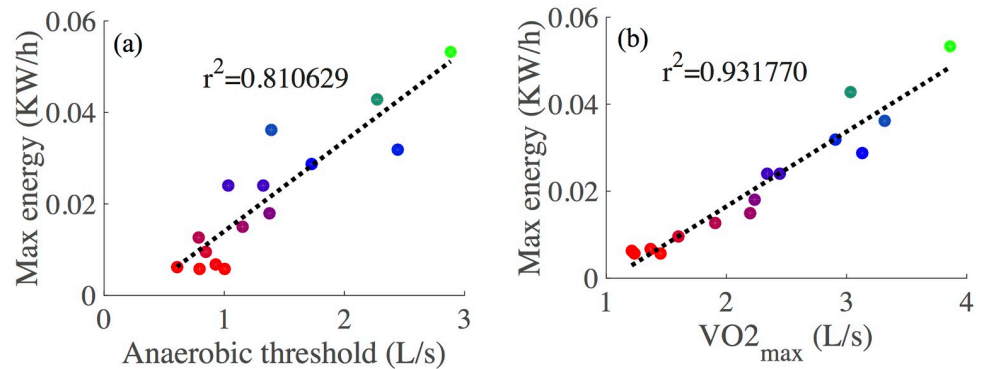
Fig 9. (a) Quantifying the relationship between the anaerobic threshold and overall test performance. Thresholds were calculated using an automated procedure based on previous methods [46] (b) $\dot{V}O_{2max}$ is the best single predictor of overall test performance.

https://doi.org/10.1371/journal.pone.0211219.g009

rate and tidal volumes, and O_2 consumption rates at a fixed ventilation rate, FVC, FEV_1, and the rate of changes in breathing rate and tidal at exercise onset, using the slopes calculated based on the analysis presented in Fig 5(b). Thus, we have an output variable that dependents on seven input variables, which we store in a vector $\mathbf{x} \in D \subset \mathbb{R}^d$, with $d = 7$. We now assume that there exists a 'true' function $f: D \rightarrow \mathbb{R}$, such that GET = $f(\mathbf{x})$.

GP emulator performance

The GP emulator was constructed using the data presented above, with GET calculated using previously described methods [46]. These data were used to train the emulator. We consider a first order polynomial regression for the mean $\mu(.)$ and the squared-exponential covariance function for $K(.,.)$. Since in this pilot study, we have a small number of participants, we use leave-one-out cross-validation mean squared error (MSE_{LOO}) to assess the accuracy of our emulator. It is defined as

$$MSE_{LOO} = \frac{1}{n}\sum_{i=1}^{n}\left(m_{-i}(\boldsymbol{x}^i) - f(\boldsymbol{x}^i)\right)^2,$$

where $m_{-i}(\boldsymbol{x}^i)$ is the prediction obtained by the GP model based on all the training data points except the $i'th$ one [47]. The corresponding prediction uncertainty is denoted by $s_{-1}^2(\boldsymbol{x}^i)$. In this work, $MSE_{LOO} = 0.1676$ which is around 12% of the global mean for GET. Fig 10 shows exact values (bullets) vs. $m_{-i}(\boldsymbol{x}^i)$ (asterisks) with the 95% confidence intervals, i.e. $m_{-i}(\boldsymbol{x}^i) \pm 1.96 s_{-i}(\boldsymbol{x}^i)$; for each patient, we use the emulator trained against the remainder of the training data to approximate the AT value for that patient, given their input variables.

The emulator has reasonable accuracy, in spite of the small number of data used to train it. In general, for high accuracy in GP emulator, the number of sample training data (corresponding to the number of patients in our case) should be around ten times larger than the number of input variables [48]. Clearly, there is a need to acquire further data points to improve the predictive capabilities of the emulator. For each patient, the 95% confidence interval around the predicted point contains the true value. However, the large variance in the GP estimates for some patients (e.g. patient 8) highlights a need to extend this study to include more data to refine estimates around these points, particularly to deal with the high variability of lung function parameters in this patient group [49].
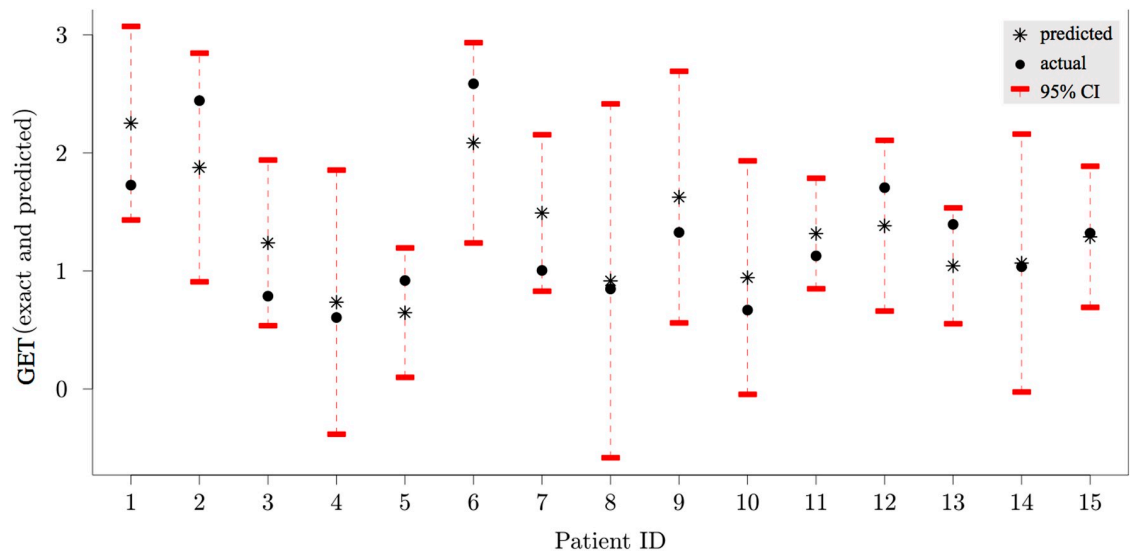
Fig 10. Predictions (asterisks) vs. exact values (bullets). Red bars show the 95% confidence intervals, $m_i(x_i) \pm 1.96\,s_i(x_i)$ around the predicted value, where m_i and s_i are GP prediction mean and standard deviation based on all but the i'th data point.

https://doi.org/10.1371/journal.pone.0211219.g010

Combining patients' data with mechanistic modelling

An alternative to the Gaussian Process-based modelling approach above would be to directly use the experimental data in a mechanistic modelling framework. Next we present an example of employing this alternative approach. During exercise, it is known that dead-space ventilation rate, $\dot{V}_D$ increases [50]. Thin et al [50], demonstrated that this increase is significantly greater in CF patients. Physiologically this is because, in CF patients, tidal volume, V_T is limited. Therefore, there must be an increase in the frequency of breathing, BF, in order to maintain gas exchange.

Fig 11(a) shows the changes in these ventilatory parameters at rest and during peak exercise for the 15 CF patients. This figure indicates that tidal volume is decreased in the CF patients as the disease severity increases. Thus, it seems reasonable to suggest that the size of increase in $\dot{V}_D$ with exercise is dependent on the severity of the disease. We can now consider this in the case of the 15 CF patients tested. For each of these patients a work rate was prescribed dependent on their previous performances (see Fig 2). We incorporate the work rate data in the mechanistic model by Timischl [51] in order to predict $\dot{V}_E$ for each individual patient (see S1 File for details on the mechanistic model). It is important to note that Timischl's original model ignores the dead-space (i.e. assumes $\dot{V}_D = 0$). Figure A in S1 File shows that generally for CF patients whose disease is less severe (those in green) $\dot{V}_E$ is captured better with no ventilatory dead-space than those whose disease is more severe. We therefore suggest incorporating the dead space by modelling $\dot{V}_D$ using a linear function of the form, $\dot{V}_D = \dot{V}_{D_B} + a \times W(t)$, where $\dot{V}_{DB}$ is the base level of $\dot{V}_D$ for each patient, a is a scaling parameter based on the severity of the disease and W(t) is the work rate. We anticipate that for those patients whose disease progression is less severe a relatively low scaling factor will be needed and for those whose condition is more severe, a much larger one. To determine the values of $\dot{V}_{DB}$ and a we use the available experimental data. It is known that, $\dot{V}_D = BF \times V_D$. As shown in Fig 11(a) the available data allows us to obtain values for BF and V_D both at rest and during exercise. We can therefore determine a value for $\dot{V}_{DB}$ as follows, $\dot{V}_{DB} = BF\ (Rest) \times V_D\ (Rest)$, and a value for a

(a)

(b)
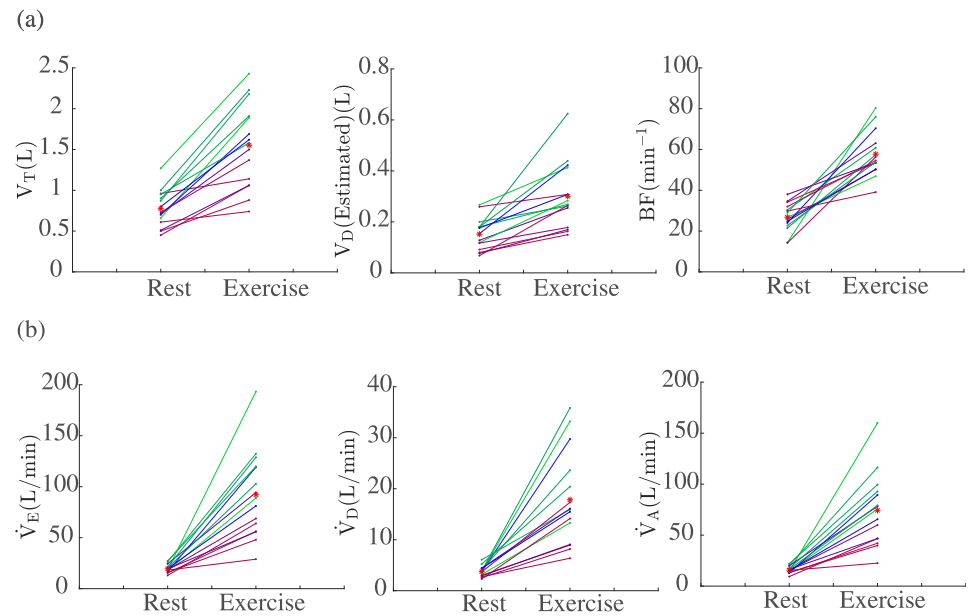
Fig 11. (a) Tidal volume (V_T), an estimate for the dead volume (V_D) and Breathing frequency BF from the data. (b) Changes in ventilatory parameters during exercise for the 15 CF patients using the proposed form of $\dot{V}_D = \dot{V}_{DB} + a \times W(t)$. In the case of $\dot{V}_E$ the solid dots represent values extracted from the data, for $\dot{V}_D$ and $\dot{V}_A$ these represent values taken from model simulations (see S1 File). In all cases the red stars represent the group mean.

https://doi.org/10.1371/journal.pone.0211219.g011

as follows, $a =$ (BF (Exercise) $\times V_D$ (Exercise)) / W (Exercise). Next, we simulate the model with this added ventilatory dead-space (model predictions are shown in Figure B in S1 File). This allows us to obtain predictions for the total ventilation ($\dot{V}_E$) for each individual patient as shown in Figure B in S1 File. The results from these simulations in turn allow us to predict alveolar ventilation, $\dot{V}_A$, for each patient as shown in Fig 11(b). In Fig 11(b) we plot $\dot{V}_E$ from the data, along with the predicted by the modified model simulations values for $\dot{V}_D$. Then using $\dot{V}_A = \dot{V}_E - \dot{V}_D$ to find $\dot{V}_A$, we can predict the changes in $\dot{V}_A$ during exercise for the 15 CF patients.

This is an example of how individual patient's data could be used in a mechanistic model framework in order to predict individual patients' characteristics, in this case associated with ventilatory parameters and hence lung function.

Discussion

Our primary long-term modelling aim is to eventually use the models to evaluate how CF impairs exercise tolerance relative to increasing ventilatory and metabolic demands. Our predictive models could also be used to evaluate therapies and their effect on exercise performance. Ultimately, we hope that this will form a series of steps to design better exercise treatment that is tailored to specific individual needs relative to patient's treatment therapies, a treatment modality that is affordable, and personalised [52]. It is interesting to note that in spite of the small number of data points used to train our GP emulator, the accuracy as computed by the leave-one-out cross-validation mean square error is high. This observation implies that the relationship between GET and our chosen input variables is 'smooth' (in that there are no large, sudden changes in GET as our input variables vary) [53]. In turn, this provides further evidence that our chosen input variables are good predictors of GET for this patient group.

The data analysis and modelling results have highlighted the dependence of CPET performance on variables associated with ventilation and metabolic rates of O_2 consumption. We believe that this is the first attempt to mathematically model the relationship between ventilatory dynamics, $\dot{V}O_2$ and performance in CPET within a group of children and adolescents with CF. Whilst it is clear that there is much work to be done in this area, we hope that this will serve as a starting point for improved modelling for CF, not only in the arena of GP emulators, but also in the domain of mechanistic modelling, which we shall describe briefly. But one of the benefits of modelling includes the ability to utilise existing data sets at a time when there are limited resources and time for a patient group who can be very sick and unable to engage in research.

Perspectives for GP improvements

In our simulator, we have used the GET as our output (dependent) variable. Another choice for this could be the performance in the CPET or $\dot{V}O_{2max}$, since these are the primary biomarkers for gauging aerobic fitness. However, the use of the GET has been shown to have high agreement with the lactate threshold (another surrogate for the AT), and related to disease severity in CF [54]. Furthermore, as reported in Fig 9(a), the GET location for a given patient correlates well with their overall performance in this test. Importantly, by constructing a predictive model to approximate the GET values for a patient, we can hope to further extend this to identify contributions of aerobic and anaerobic pathways in supplying ATP to meet the demand imposed during the exercise test.

The initial exploration of results highlighted that both ventilation parameters and metabolic rates of O_2 consumption were the primary factors influencing test performance. It is clear that $\dot{V}O_2$ should play a significant role in determining the GET location, since it is a proxy for oxidative phosphorylation which is the main pathway for ATP synthesis in steady state exercise. As a measure of oxygen uptake efficiency in our model, we use the oxygen consumption rate at a fixed total ventilation rate (that being 0.822 L·s^{-1}) as an input (independent) variable for each patient.

There are a number of ventilatory input variables incorporated in our simulator. Given their potential importance as clinical biomarkers, highlighting the limitations of lung capacity and function, we include FVC and FEV_1 as input (independent) variables. During the aerobic exercise test, participants spend three minutes cycling at a minimal work rate, over which we quantify their baseline breathing frequency and baseline tidal volume by taking the means of these variables over this period. To capture the dynamics response associated with the exercise, the rates of change of breathing frequency and tidal volume are calculated, based on the fits obtained in Fig 5(b). The rates of change of these ventilatory variables indicate how individual participants respond to changes in work rate and were shown in Fig 7(b) to discriminate between participant performances. Moreover, differences in rates of change of breathing frequency and tidal volume have previously been shown to be significantly different between control groups and CF groups [50], suggesting that these are potentially key biomarkers for assessing aerobic fitness in patients with CF.

At present, the GP model is conditioned on specific data points for each patient. An improvement to the GP could be made by instead conditioning with respect to distributions. Given that repeated tests are often performed for the same individual, so that multiple sample points are provided for each participant, we can consider a fit to a probability distribution capturing the variability in the identified variables. This approach has advantages compared to standard GP models, such as avoiding problems associated with over fitting and regularisation

(which is important for the inverting ill-conditioned covariance matrices that often arise during the application of [1, 2, 55]).

Perspectives for mechanistic modelling

In order to better understand and characterise the difference between performances, it would be extremely informative to construct and simulate a mechanistic mathematical model, based upon on an ordinary differential equation (ODE) framework, describing the relationship between the cardiopulmonary system and the metabolic dynamics of skeletal muscle. By describing the relationships between different organ-level systems, the model would be able to identify the patient-specific rate-limiting factors defining aerobic fitness. Moreover, analysis of the model could be used to suggest treatment strategies to improve these factors and thus predict how patients will improve under such regimes.

At the individual organ level, there are a plethora of models describing individual dynamics of the level of the heart [56–59], lung [60–63] and systemic metabolic demand [64–70]. There also exist a number of models describing such interactions between cardiopulmonary and metabolic systems [51, 71–74] in a variety of settings, including heart failure and mechanical ventilation. A core feature in all of these models is the nonlinear interactions between the constituent model compartments that encompass the distinct tissues. An important consequence of this is that the model must be studied as whole, in an integrated fashion, to truly understand the body's response to exercise.

With respect to the present question, there are a number of limitations of the existing modelling approaches. Most significantly, none have been designed with either an adolescent, or a CF patient group in mind, and the nuances of these patient groups will have to be factored into to any model development. In particular, these models have relatively simple, empirical models to describe changes in ventilation, which may not capture well the breathing dynamics of our patient group. Moreover, to the best of our knowledge, no model considers the changes in ventilation separated into breathing frequency and depth that have been shown by us and others [50] to be critical to overall test performance.

In our analysis, we have demonstrated that the GET location is a critical factor in determining overall patient aerobic fitness. Many of the mathematical exercise models describe only steady-state exercise, in which aerobic pathways meet most of the ATP demand [51, 71–74]. As such, these models are inadequate to capture the dynamics we describe here. Another common topic of study is the dynamic response at exercise onset, which again, does not meet the current need to describe the GET crossing point [75–78].

Of the mathematical models that describe the contribution of anaerobic pathways to ATP production, some assume that the shortfall in meeting ATP demand via oxidative phosphorylation is met entirely by anaerobic pathways [79], yet this is clearly not so, since ATP levels in skeletal muscle post-exercise may be up to 30% lower than pre-exercise values [80]. Mathematical models that factor in fatigue brought about by anaerobic metabolism are generally phenomenological in nature, and it is difficult to quantify these models against real patient data [79, 81, 82]. Moreover, these models mostly fall outside the arena of ODE-based modelling and so dynamical properties are difficult to infer from them.

Developing a mathematical model to describe the integrated behaviour of all of the relevant organs, whilst remaining biophysically plausible, but without requiring excessive or invasive parameterisation, is a difficult task. The proposed model should include descriptions of the cardiovascular system, the ventilatory system and simple models of metabolism at the tissue level. Specifically, dynamic variables should include alveolar, arterial, venous and tissue level partial pressures/concentrations of O_2 and CO_2, cardiac output, ventilation and metabolic
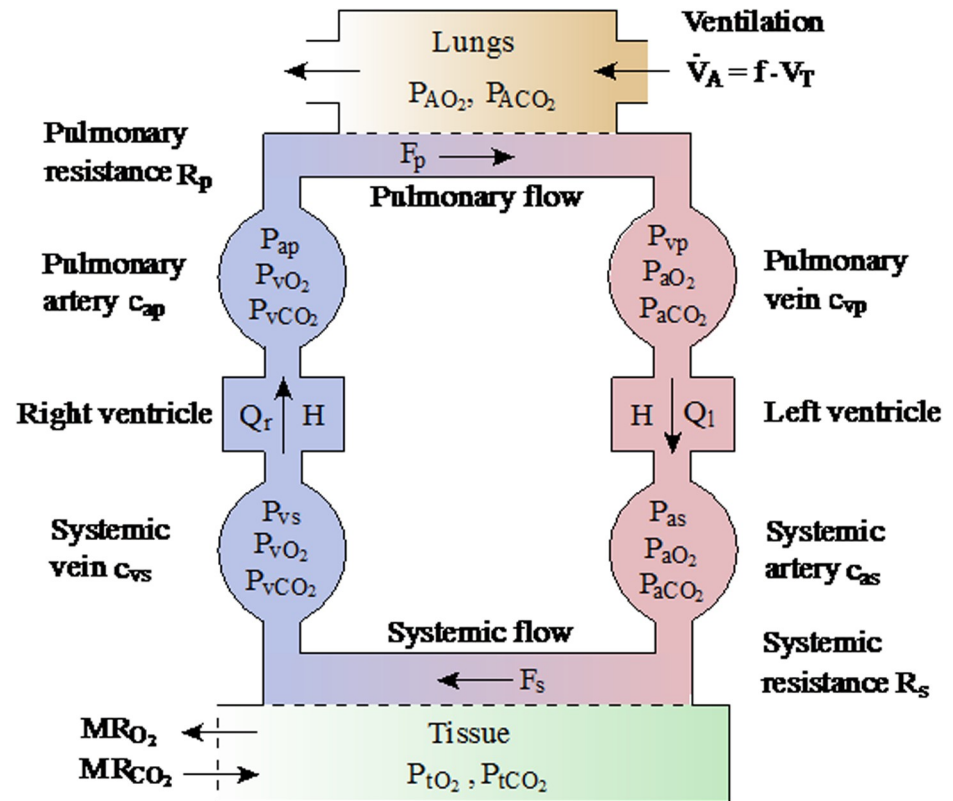
Fig 12. Schematic of the variables and processes in the proposed ODE-based mathematical model. Adapted from Timischl [51] and Batzel *et al.* [86].

rates oxygen utilisation and CO_2 production. Partial alveolar gas pressures can be linked to data collected during the test, and the work rate can then be provided as inputs to the model. Note that these variables are similar to those included in previously defined models [51, 71–74] and the aim is to extend these to describe the dynamics observed in patients with CF. The proposed model schematic is displayed in Fig 12.

Of critical importance to the overall model construction is the development of a simple, yet realistic model of cellular metabolism, to overcome the issues discussed earlier. The model should respect the different metabolic processes that occur in the muscle tissue, in particular: glycolysis, phosphocreatine breakdown and synthesis and oxidative phosphorylation, in a simplistic fashion that is amenable to being fit to CPET data. Whilst there are models that describe the biochemical reactions associated with these processes, and importantly, their stoichiometry [65, 68, 70, 73, 83], quantifying their associated rate constants *in vivo* is a near-impossible task, and so efforts must be made to develop a model that incorporates the relevant metabolic dynamics whilst being simple enough to be fit to data.

With knowledge of the integrated system, attempts can also be made to describe other important exercise-based processes, such as lactate buffering and recycling (as a fuel source) [68, 84, 85] and the overall muscle fatigue brought about by the combination of all of these factors. Only by systematically exploring the dependence of aerobic fitness of all of the factors described in this section can we begin to understand the system in an integrated fashion.

Limitations

A limitation with the current study is the utilisation of a relatively small sample size, and this is most likely contributing towards aforementioned errors in prediction. Future studies should seek to utilise CPET collected annually in CF centres, to develop larger, multi-centre, samples whereby a uniform exercise protocol is utilised. Given that utilisation of CPET is now recommended and endorsed for regular use by international medical societies [87], and individual CF centres are reporting upon experiences of using CPET [88], large-scale utilisation of such data is a feasible target.

The findings presented here are derived from a smaller sample, and therefore the models presented are only preliminary models of this patient cohort; however, our study provides a unique examination into the aerobic and anaerobic signatures of individual patients with CF in response to progressive exercise.

As models improve, as well as the quality of fits to data (through increased sample sizes), these can be used in a prognostic setting to predict potential improvements in aerobic fitness that may arise due to therapeutic intervention. Moreover, with proper mechanistic modelling of the primary organs affected in CF, there exists the potential to optimise treatment for this patient group by identifying the limiting factors of aerobic fitness. Finally, whilst this study provides an insight into metabolic process during exercise, future research and models must account for additional variables predictive of function and mortality (e.g. genotype, body composition, pancreatic sufficiency, infection status, exacerbations [89, 90]) and co-morbidities (CF-related diabetes [91], pulmonary arterial hypertension [92]) existent within CF, notably those that may affect exercise tolerance.

Conclusions

In conclusion, our proposed stimulator allows for the reproduction of the physiological observations of ventilation and metabolic rates of O_2 consumption acquired during a CPET in relation to people with CF. The modelling framework was able to successfully replicate the relationship between ventilatory dynamics, lung capacity and function and performance in CPET within a group of children and adolescents with CF. In particular, by using the Gaussian processes (GP) the GET, a well-known physiological threshold marker of exercise intensity, as well as other important measurements such as breathing frequency and tidal volume, at the individual patient level were accommodated into the model. The stimulator has the potential to be suitable for future applications of the investigations of drug therapies or other physical interventions on exercise performance.

Supporting information

S1 File. Cardiopulmonary responses to maximal aerobic exercise in patients with cystic fibrosis. Figure A: Model simulations of $\dot{V}_E$ assuming $V_D = 0$, compared to the recorded data value of $\dot{V}_E$ obtained during the CPET. The thin red line gives $\dot{V}_E$ as measured from the data and the bold coloured line gives $\dot{V}_E$ as simulated by the model. The colour given to each patient is the same as given in Fig 2 in the main text of the paper. **Figure B**: Model simulations of $\dot{V}_E$ using the proposed linear model of V_D, compared to the recorded data value of $\dot{V}_E$ obtained during the CPET. The thin red line gives $\dot{V}_E$ as measured from the data and the bold coloured line gives $\dot{V}_E$ as simulated by the model. The colour given to each patient is the same as given in in Fig 2 in the main text of the paper. Fitted parameter values can be found below.

Table A: Parameter values. **Table B**: Calculated values of $\dot{V}_{D_B}$ and a for the linear model of ventilatory dead-space. **Table C**: Definitions of model parameters.
(PDF)

Author Contributions

Conceptualization: Craig A. Williams.

Data curation: Craig A. Williams, Owen W. Tomlinson.

Formal analysis: Craig A. Williams, Kyle C. A. Wedgwood, Hossein Mohammadi, Katie Prouse, Owen W. Tomlinson, Krasimira Tsaneva-Atanasova.

Investigation: Craig A. Williams.

Methodology: Craig A. Williams, Kyle C. A. Wedgwood, Krasimira Tsaneva-Atanasova.

Project administration: Craig A. Williams, Owen W. Tomlinson.

Supervision: Craig A. Williams.

Writing – original draft: Craig A. Williams, Kyle C. A. Wedgwood, Krasimira Tsaneva-Atanasova.

Writing – review & editing: Craig A. Williams, Kyle C. A. Wedgwood, Hossein Mohammadi, Katie Prouse, Owen W. Tomlinson, Krasimira Tsaneva-Atanasova.

References

1. Cystic Fibrosis Trust. UK cystic fibrosis registry 2015 annual data report. London, UK; 2016.

2. Elborn JS. Cystic fibrosis. The Lancet. 2016; 388(10059):2519–31.

3. Pérez M, Groeneveld IF, Santana-Sosa E, Fiuza-Luces C, Gonzalez-Saiz L, Villa-Asensi JR, et al. Aerobic fitness is associated with lower risk of hospitalization in children with cystic fibrosis. Pediatric Pulmonology. 2014; 49(7):641–9. https://doi.org/10.1002/ppul.22878 PMID: 24019231

4. Andréasson B, Jonson B, Kornfält R, Nordmark EVA, Sandström S. Long-term Effects of Physical Exercise on Working Capacity and Pulmonary Function in Cystic Fibrosis. Acta paediatrica. 1987; 76(1):70–5.

5. Orenstein DM. Exercise conditioning and cardiopulmonary fitness in cystic fibrosis. The effects of a three-month supervised running program. Chest. 1981; 80(4):392. PMID: 7023863

6. Salh W, Bilton D, Dodd M, Webb AK. Effect of exercise and physiotherapy in aiding sputum expectoration in adults with cystic fibrosis. Thorax. 1989; 44(12):1006–8. PMID: 2617441

7. Schneiderman-Walker J, Pollock SL, Corey M, Wilkes DD, Canny GJ, Pedder L, et al. A randomized controlled trial of a 3-year home exercise program in cystic fibrosis. The Journal of Pediatrics. 2000; 136(3):304–10. https://doi.org/10.1067/mpd.2000.103408 PMID: 10700685

8. Nixon PA, Orenstein DM, Kelsey SF, Doershuk CF. The prognostic value of exercise testing in patients with cystic fibrosis. New England Journal of Medicine. 1992; 327(25):1785–8. https://doi.org/10.1056/NEJM199212173272504 PMID: 1435933

9. Pianosi P, LeBlanc J, Almudevar A. Peak oxygen uptake and mortality in children with cystic fibrosis. Thorax. 2005; 60(1):50–4. https://doi.org/10.1136/thx.2003.008102 PMID: 15618583

10. Cystic Fibrosis Trust. Standards of care and good clinical practice for the physiotherapy management of cystic fibrosis. London, UK; 2017.

11. Stevens D, Oades PJ, Armstrong N, Williams CA. Exercise metabolism during moderate-intensity exercise in children with cystic fibrosis following heavy-intensity exercise. Applied Physiology, Nutrition, and Metabolism. 2011; 36(6):920–7. https://doi.org/10.1139/h11-117 PMID: 22087846

12. Williams CA, Stevens D. Physical activity and exercise training in young people with cystic fibrosis: Current recommendations and evidence. Journal of Sport and Health Science. 2013; 2(1):39–46.

13. Angelis A, Kanavos P, Lopez-Bastida J, Linertova R, Nicod E, Serrano-Aguilar P, et al. Social and economic costs and health-related quality of life in non-institutionalised patients with cystic fibrosis in the United Kingdom. BMC Health Services Research. 2015; 15(1):428.

14. Wheatley CM, Wilkins BW, Snyder EM. Exercise is medicine in cystic fibrosis. Exercise and Sport Sciences Reviews. 2011; 39(3):155–60. https://doi.org/10.1097/JES.0b013e3182172a5a PMID: 21383626

15. Stevens D, Oades PJ, Armstrong N, Williams CA. A survey of exercise testing and training in UK cystic fibrosis clinics. Journal of Cystic Fibrosis. 2010; 9(5):302–6. https://doi.org/10.1016/j.jcf.2010.03.004 PMID: 20359963

16. Goodson HV, Gregoretti IV. Using Computational Modeling to Understand Microtubule Dynamics. Methods in Cell Biology. 2010; 95(10):175–88.

17. Hardman JG, Ross JJ. Modelling: a core technique in anaesthesia and critical care research. British Journal of Anaesthesia. 2006; 97(5):589–92. https://doi.org/10.1093/bja/ael272 PMID: 17032663

18. Holford N, Ma SC, Ploeger BA. Clinical trial simulation: a review. Clinical pharmacology and therapeutics. 2010; 88(2):166–82. https://doi.org/10.1038/clpt.2010.114 PMID: 20613720

19. Kapellos GE, Alexiou TS, Payatakes AC. Theoretical modeling of fluid flow in cellular biological media: an overview. Mathematical biosciences. 2010; 225(2):83–93. https://doi.org/10.1016/j.mbs.2010.03.003 PMID: 20303992

20. Seethala RR, Esposito EC, Abella BS. Approaches to improving cardiac arrest resuscitation performance. Current opinion in critical care. 2010; 16(3):196–202. https://doi.org/10.1097/MCC.0b013e328338c121 PMID: 20305553

21. McCahon RA, Columb MO, Mahajan RP, Hardman JG. Validation and application of a high-fidelity, computational model of acute respiratory distress syndrome to the examination of the indices of oxygenation at constant lung-state. British Journal of Anaesthesia. 2008; 101(3):358–65. https://doi.org/10.1093/bja/aen181 PMID: 18567677

22. Mintun MA, Lundstrom BN, Snyder AZ, Vlassenko AG, Shulman GL, Raichle ME. Blood flow and oxygen delivery to human brain during functional activity: Theoretical modeling and experimental data. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98(12):6859–64. https://doi.org/10.1073/pnas.111164398 PMID: 11381119

23. Noble D. Modeling the heart–from genes to cells to the whole organ. Science. 2002; 295(5560):1678–82. https://doi.org/10.1126/science.1069881

24. Moppett IK, Hardman JG. Development and Validation of an Integrated Computational Model of Cerebral Blood Flow and Oxygenation. Anesthesia & Analgesia. 2007; 105(4):1094–103.

25. Murkin JM, Arango M. Near-infrared spectroscopy as an index of brain and tissue oxygenation. British Journal of Anaesthesia. 2009; 103(Supplement 1):i3–i13.

26. Hermans MP, Brotons C, Elisaf M, Michel G, Muls E, Nobels F. Optimal type 2 diabetes mellitus management: the randomised controlled OPTIMISE benchmarking study: baseline results from six European countries. European Journal of Preventive Cardiology. 2013; 20(6):1095–105. https://doi.org/10.1177/2047487312449414 PMID: 22605788

27. Saynor ZL, Barker AR, Oades PJ, Williams CA. A protocol to determine valid VO_{2max} in young cystic fibrosis patients. Journal of Science and Medicine in Sport. 2013; 16(6):539–44.

28. Saynor ZL, Barker AR, Oades PJ, Williams CA. Reproducibility of maximal cardiopulmonary exercise testing for young cystic fibrosis patients. Journal of Cystic Fibrosis. 2013; 12(6):644–50. https://doi.org/10.1016/j.jcf.2013.04.012 PMID: 23726255

29. Hulzebos HJ, Werkman MS, van Brussel M, Takken T. Towards an individualized protocol for workload increments in cardiopulmonary exercise testing in children and adolescents with cystic fibrosis. Journal of Cystic Fibrosis. 2012; 11(6):550–4. https://doi.org/10.1016/j.jcf.2012.05.004 PMID: 22704761

30. Barker AR, Williams CA, Jones AM, Armstrong N. Establishing maximal oxygen uptake in young people during a ramp cycle test to exhaustion. British Journal of Sports Medicine. 2011; 45(6):498–503. https://doi.org/10.1136/bjsm.2009.063180 PMID: 19679577

31. Saynor ZL, Barker AR, Oades PJ, Williams CA. Impaired aerobic function in patients with cystic fibrosis during ramp exercise. Medicine and Science in Sports and Exercise. 2014; 46(12):2271–778. https://doi.org/10.1249/MSS.0000000000000369 PMID: 24781889

32. Sacks J, Welch WJ, Mitchell TJ, Wynn HP. Design and analysis of computer experiments. Statistical Science. 1989:409–23.

33. Kleijnen JP. Kriging metamodeling in simulation: A review. European Journal of Operational Research. 2009; 192(3):707–16.

34. Jin R, Chen W, Simpson TW. Comparative studies of metamodelling techniques under multiple modelling criteria. Structural and Multidisciplinary Optimization. 2001; 23(1):1–13.

35. Forrester AI, Keane AJ. Recent advances in surrogate-based optimization. Progress in Aerospace Sciences. 2009; 45(1):50–79.

36. Chen VCP, Tsui KL, Barton RR, Meckesheimer M. A review on design, modeling and applications of computer experiments. IIE Transactions. 2006; 38(4):273–91.

37. Schwaighofer A, Grigoras M, Tresp V, Hoffmann C, editors. GPPS: A Gaussian process positioning system for cellular networks. Advances in Neural Information Processing Systems; 2004.

38. Bailer-Jones CAL, Sabin TJ, MacKay DJC, Withers PJ, editors. Prediction of deformed and annealed microstructures using Bayesian neural networks and Gaussian processes. Proceedings of the Australasia Pacific Forum on Intelligent Processing and Manufacturing of Materials; 1997.

39. Kalaitzis AA, Lawrence ND. A simple approach to ranking differentially expressed gene expression time courses through Gaussian process regression. BMC Bioinformatics. 2011; 12(1):180.

40. Lawrence ND, Sanguinetti G, Rattray M. Modelling transcriptional regulation using Gaussian processes. Advances in Neural Information Processing Systems. 2007; 19:785.

41. Swain PS, Stevenson K, Leary A, Montano-Gutierrez LF, Clark IB, Vogel J, et al. Inferring time derivatives including cell growth rates using Gaussian processes. Nature Communications. 2016; 7.

42. Rasmussen C, Williams C. Gaussian processes for machine learning (Adaptive computation and machine learning series) Cambridge, MA: The MIT Press; 2005.

43. Taylor-Robinson D, Whitehead M, Diderichsen F, Olesen HV, Pressler T, Smyth RL, et al. Understanding the natural progression in% FEV1 decline in patients with cystic fibrosis: a longitudinal study. Thorax. 2012; 67(10):860–6. https://doi.org/10.1136/thoraxjnl-2011-200953 PMID: 22555277

44. Vilozni D, Bentur L, Efrati O, Minuskin T, Barak A, Szeinberg A, et al. Spirometry in early childhood in cystic fibrosis patients. Chest. 2007; 131(2):356–61. https://doi.org/10.1378/chest.06-1351 PMID: 17296633

45. Kerem E, Reisman J, Corey M, Canny GJ, Levison H. Prediction of mortality in patients with cystic fibrosis. New England Journal of Medicine. 1992; 326(18):1187–91. https://doi.org/10.1056/NEJM199204303261804 PMID: 1285737

46. Beaver WL, Wasserman K, Whipp BJ. A new method for detecting anaerobic threshold by gas exchange. Journal of Applied Physiology. 1986; 60(6):2020–7. https://doi.org/10.1152/jappl.1986.60.6.2020 PMID: 3087938

47. Jin R, Chen W, Sudjianto A, editors. On sequential sampling for global metamodelling in engineering design. Proceedings of DETC'02 ASME 2002 Design Engineering Technical Conferences And Computers and Information in Engineering Conference; 2002.

48. Jones DR, Schonlau M, Welch WJ. Efficient global optimization of expensive black-box functions. Journal of Global Optimization. 1998; 13(4):455–92.

49. Cooper PJ, Robertson CF, Hudson IL, Phelan PD. Variability of pulmonary function tests in cystic fibrosis. Pediatric Pulmonology. 1990; 8(1):16–22. PMID: 2405342

50. Thin AG, Dodd JD, Gallagher CG, Fitzgerald MX, McLoughlin P. Effect of respiratory rate on airway deadspace ventilation during exercise in cystic fibrosis. Respiratory Medicine. 2004; 98(11):1063–70. PMID: 15526806

51. Timischl S. A Global Model for the Cardiovascular and Respiratory System: University of Graz; 1998.

52. Balfour-Lynn IM. Personalised medicine in cystic fibrosis is unaffordable. Paediatric Respiratory Reviews. 2014; 15 Suppl 1:2–5.

53. Forrester AIJ, Sobester A, Keane AJ, editors. Multi-fidelity optimisation via surrogate modelling. Proceedings: Mathematical, Physical and Engineering Sciences; 2007.

54. Thin AG, Linnane SJ, McKone EF, Freaney R, FitzGerald MX, Gallagher CG, et al. Use of the gas exchange threshold to noninvasively determine the lactate threshold in patients with cystic fibrosis. Chest. 2002; 121(6):1761–70. PMID: 12065336

55. Titsias MK, editor Variational Learning of Inducing Variables in Sparse Gaussian Processes. AISTATS; 2009.

56. Mazzoleni MJ, Battaglini CL, Martin KJ, Coffman EM, Mann BP. Modeling and predicting heart rate dynamics across a broad range of transient exercise intensities during cycling. Sports Engineering. 2016; 19(2):117–27.

57. Stirling JR, Zakynthinaki M, Refoyo I, Sampedro J. A Model of Heart Rate Kinetics in Response to Exercise. Journal of Nonlinear Mathematical Physics. 2008; 15(sup3):426–36.

58. Ursino M. Interaction between carotid baroregulation and the pulsating heart: a mathematical model. American Journal of Physiology-Heart and Circulatory Physiology. 1998; 275(5):H1733–H47.

59. Zakynthinaki MS. Modelling heart rate kinetics. PLoS One. 2015; 10(4):1–26.

60. Gutierrez G. A mathematical model of tissue–blood carbon dioxide exchange during hypoxia. American Journal of Respiratory and Critical Care Medicine. 2004; 169(4):525–33. https://doi.org/10.1164/rccm.200305-702OC

61. Bates JHT. Lung mechanics: an inverse modeling approach. Cambridge: Cambridge University Press; 2009.

62. Reynolds A, Ermentrout GB, Clermont G. A mathematical model of pulmonary gas exchange under inflammatory stress. Journal of Theoretical Biology. 2010; 264(2):161–73. https://doi.org/10.1016/j.jtbi.2010.01.011 PMID: 20083125

63. Crooke P, Hotchkiss J, Marini J. Linear and nonlinear mathematical models for noninvasive ventilation. Mathematical and Computer Modelling. 2002; 35(11–12):1297–313.

64. Kek K, Miyakawa T, Yoneyama S, Kudo N, Yamamoto K, editors. Simulation of exercise-dependent difference in metabolism with a mathematical model for analyses of measurements using near-infrared spectroscopy. Engineering in Medicine and Biology Society, 2006 EMBS'06 28th Annual International Conference of the IEEE; 2006: IEEE.

65. Krauss M, Schaller S, Borchers S, Findeisen R, Lippert J, Kuepfer L. Integrating cellular metabolism into a multiscale whole-body model. PLoS Computational Biology. 2012; 8(10):e1002750. https://doi.org/10.1371/journal.pcbi.1002750 PMID: 23133351

66. Li Y, Lai N, Kirwan JP, Saidel GM. Computational model of cellular metabolic dynamics in skeletal muscle fibers during moderate intensity exercise. Cellular and Molecular Bioengineering. 2012; 5(1):92–112. PMID: 22942911

67. Lai N, Camesasca M, Saidel GM, Dash RK, Cabrera ME. Linking pulmonary oxygen uptake, muscle oxygen utilization and cellular metabolism during exercise. Annals of Biomedical Engineering. 2007; 35(6):956–69. https://doi.org/10.1007/s10439-007-9271-4 PMID: 17380394

68. Cloutier M, Wellstead P. The control systems structures of energy metabolism. Journal of The Royal Society Interface. 2009:rsif20090371.

69. Dash RK, Dibella Ja, M E. A computational model of skeletal muscle metabolism linking cellular adaptations induced by altered loading states to metabolic responses during exercise. Biomedical Engineering Online. 2007; 6:14. https://doi.org/10.1186/1475-925X-6-14 PMID: 17448235

70. Kim J, Saidel GM, Cabrera ME. Multi-scale computational model of fuel homeostasis during exercise: effect of hormonal control. Annals of Biomedical Engineering. 2007; 35(1):69–90. https://doi.org/10.1007/s10439-006-9201-x PMID: 17111212

71. Fresiello L, Meyns B, Di Molfetta A, Ferrari G. A Model of the Cardiorespiratory Response to Aerobic Exercise in Healthy and Heart Failure Conditions. Frontiers in Physiology. 2016; 7.

72. Fresiello L, Ferrari G, Di Molfetta A, Zieliński K, Tzallas A, Jacobs S, et al. A cardiovascular simulator tailored for training and clinical uses. Journal of Biomedical Informatics. 2015; 57:100–12. https://doi.org/10.1016/j.jbi.2015.07.004 PMID: 26173037

73. D'Angelo C. Multiscale modelling of metabolism and transport phenomena in living tissues: École Polytechnique Fédérale de Lausanne; 2007.

74. Hardman J, Bedforth N, Ahmed A, Mahajan R, Aitkenhead A. A physiology simulator: validation of its respiratory components and its ability to predict the patient's response to changes in mechanical ventilation. British Journal of Anaesthesia. 1998; 81(3):327–32. PMID: 9861113

75. Lai N, Saidel GM, Grassi B, Gladden LB, Cabrera ME. Model of oxygen transport and metabolism predicts effect of hyperoxia on canine muscle oxygen uptake dynamics. Journal of Applied Physiology. 2007; 103(4):1366–78. https://doi.org/10.1152/japplphysiol.00489.2007 PMID: 17600157

76. Benson AP, Grassi B, Rossiter HB. A validated model of oxygen uptake and circulatory dynamic interactions at exercise onset in humans. Journal of Applied Physiology. 2013; 115(5):743–55. https://doi.org/10.1152/japplphysiol.00184.2013 PMID: 23766506

77. Bell C, Paterson DH, Kowalchuk JM, Padilla J, Cunningham DA. A Comparison of Modelling Techniques used to Characterise Oxygen Uptake Kinetics During the on-Transient of Exercise. Experimental Physiology. 2001; 86(5):667–76. PMID: 11571496

78. Elstad M, Toska K, Walløe L. Model simulations of cardiovascular changes at the onset of moderate exercise in humans. The Journal of Physiology. 2002; 543(2):719–28.

79. Ng LJ, Sih BL, Stuhmiller JH. An integrated exercise response and muscle fatigue model for performance decrement estimates of workloads in oxygen-limiting environments. European Journal of Applied Physiology. 2012; 112(4):1229–49. https://doi.org/10.1007/s00421-011-2062-5 PMID: 21769737

80. Walter G, Vandenborne K, Elliott M, Leigh JS. In vivo ATP synthesis rates in single human muscles during high intensity exercise. The Journal of Physiology. 1999; 519(3):901–10.

81. Eriksson A, Holmberg H-C, Westerblad H. A numerical model for fatigue effects in whole-body human exercise. Mathematical and Computer Modelling of Dynamical Systems. 2016; 22(1):21–38.

82. Liu JZ, Brown RW, Yue GH. A dynamical model of muscle activation, fatigue, and recovery. Biophysical Journal. 2002; 82(5):2344–59. https://doi.org/10.1016/S0006-3495(02)75580-X PMID: 11964225

83. Dash RK, Li Y, Kim J, Saidel GM, Cabrera ME. Modeling cellular metabolism and energetics in skeletal muscle: large-scale parameter estimation and sensitivity analysis. IEEE Transactions on Biomedical Engineering. 2008; 55(4):1298–318. https://doi.org/10.1109/TBME.2007.913422 PMID: 18390321

84. Moxnes JF, Sandbakk Ø. The kinetics of lactate production and removal during whole-body exercise. Theoretical Biology and Medical Modelling. 2012; 9(1):7.

85. Péronnet F, Aguilaniu B. Lactic acid buffering, nonmetabolic CO2 and exercise hyperventilation: a critical reappraisal. Respiratory Physiology & Neurobiology. 2006; 150(1):4–18.

86. Batzel JJ, Kappel F, Timischl-Teschl S. A cardiovascular-respiratory control system model including state delay with application to congestive heart failure in humans. Journal of mathematical biology. 2005; 50(3):293–335. https://doi.org/10.1007/s00285-004-0293-3 PMID: 15480669

87. Hebestreit H, Arets HG, Aurora P, Boas S, Cerny F, Hulzebos EH, et al. Statement on exercise testing in cystic fibrosis. Respiration. 2015; 90(4):332–51. https://doi.org/10.1159/000439057 PMID: 26352941

88. Weir E, Burns PD, Devenny A, Young D, Paton JY. Cardiopulmonary exercise testing in children with cystic fibrosis: one centre's experience. Archives of disease in childhood. 2017; 102(5):440–4. https://doi.org/10.1136/archdischild-2016-310651 PMID: 27919886

89. Liou TG, Adler FR, Fitzsimmons SC, Cahill BC, Hibbs JR, Marshall BC. Predictive 5-Year survivorship model of cystic fibrosis. American Journal of Epidemiology. 2001; 153(4):345–52. PMID: 11207152

90. Aaron SD, Stephenson AL, Cameron DW, Whitmore GA. A statistical model to predict one-year risk of death in patients with cystic fibrosis. Journal of Clinical Epidemiology. 2015; 68(11):1336–45. https://doi.org/10.1016/j.jclinepi.2014.12.010 PMID: 25655532

91. Middleton PG, Matson AG, Robinson PD, Jane Holmes-Walker D, Katz T, Hameed S. Cystic Fibrosis Related Diabetes: Potential pitfalls in the transition from paediatric to adult care. Paediatric Respiratory Reviews. 2014; 15(3):281–4. https://doi.org/10.1016/j.prrv.2014.05.003 PMID: 24973264

92. Hayes D Jr., Tobias JD, Mansour HM, Kirkby S, McCoy KS, Daniels CJ, et al. Pulmonary hypertension in cystic fibrosis with advanced lung disease. American Journal of Respiratory and Critical Care Medicine. 2014; 190(8):898–905. https://doi.org/10.1164/rccm.201407-1382OC PMID: 25222938