**Supplement 1. Development of artificial neural network (ANN) and logistic regression models**

*Development of the logistic regression model*

After data was collected, the predictors for every subject were placed in a table, where every row corresponds to one subject, and the corresponding outcomes of all subjects were placed in the final row of that table. Then the coefficient for every predictor is computed using the least square method as implemented in the statistics and machine learning toolbox in Matlab.

This method yields the best linear model that fits the data. By best we mean the model that minimizes the square error. Logistic regression simply uses the output *y* of the linear regression in the following equation to output the prediction p:

$$p=\frac{1}{1+e^{-y}}$$

If p < 0.5, the prediction is negative otherwise the prediction is positive.

*Development of ANN models*

The training of the network was done using the Levenberg-Marquardt back-propagation algorithm. This algorithm finds the weights that minimize the error using a variation of Newton’s method for minimizing functions [1]. This algorithm was chosen because it is the fastest neural networks training algorithm for moderate size networks [2] as is the case in this study. The activation functions used in each NN layer is the rectified linear unit (ReLu). This activation function is being used extensively in the current literature for continuous valued data like strain values. The training was done using the Matlab software. “m” scripts were used to implement the whole architecture and used the Levenberg-Marquardt back propagation algorithm as implemented in the Matlab Neural Network toolbox. On Matlab the data was split into a derivation cohort used to derive the prediction model and a validation cohort used to measure the performance of the derived model.

In this study the most accurate full feature model (using all significantly predictive variables) that yielded the best results has an input layer of 2 neurons that are connected to all the observed independent variables. The hidden layer consists of 2 neurons and the output layer of 1 neuron.

The most accurate strain based model (using variables in bold in table 2) has an input layer of 2 neurons that are connected to all the observed independent variables. The hidden layer consists of 2 neurons and the output layer of 1 neuron.

1. Scales LE. Introduction to non-linear optimization MacMillan, City and Guilds. 1985.

2. Hagan MT, H.B. Demuth, M.H. Beale. Neural Network Design. Boston, MA, USA: PWS Publishing Co. 1996.