

# HPC vital to predicting patient response from big data clinical trials



BAST Inc. combined HPC Midlands massively parallel computing environment with open-source software to quickly assess large clinical trial datasets and develop models that predict patient responses to novel medical treatment.

## Challenge

With the advance of high throughput methods, ever increasing amounts of data are collected during the recruitment of patients to clinical trials. The Loughborough University Science and Enterprise Parks based company BAST Inc. offers advanced data analysis services in order to relate such recorded data to clinical responses and identify biomarkers that indicate if a patient will respond to a novel treatment.

When dealing with such data sets often the challenge is that the number of descriptor variables exceeds the number of available data points. To identify predictive and statistically significant biomarker signatures, computationally heavy approaches have to be employed that are easily beyond the limits of what is typically available to small enterprises.

## Solution

Together with their main client, a large, multinational pharmaceutical company, BAST used the HPC Midlands cluster, Hera, to compare the suitability of several machine learning approaches to identify potential biomarker signatures. With Hera's large core capability and low-latency interconnect system BAST was able to meet their stakeholders severe time constraints in establishing which method is most suitable for the task at hand. Being able to analyse the data quickly resulted in the identification of several potential biomarker signatures predicting clinical response and selecting the most promising candidate model.

The project's large datasets were ported to Hera through a high-speed I/O network, JANET, while BAST staff members were able to work from their own desktops and interact with their software application, R, the same as they do on smaller projects.

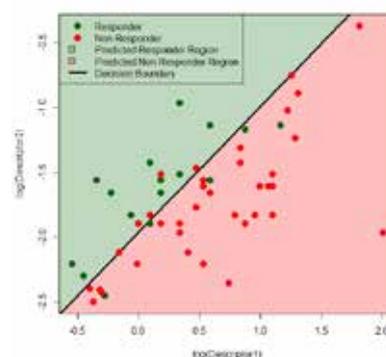
## Impact

A hypothesis was generated as to which biomarker signature is most likely to predict whether a patient will respond to a specific treatment. If validated by a follow up study, the stakeholder would be in the possession of a predictive model with which only patients likely to respond are selected for treatment. This strategy would greatly increase the probability of success of the complete drug development programme, and potentially nonresponsive patients could be assigned to alternative treatments.



**"For a small company to be competitive in the analysis of increasingly large clinical trial data sets it is absolutely vital to have on-demand access on short notice to a high performance computing facility as the one offered by HPC Midlands."**

Garrit Jentsch, Senior Scientist, BAST Inc.



Without leaving their office BAST was able to parallel process large clinical trial datasets to stratify patients faster and better predict their responses to treatment.

HPC Midlands is the East Midlands' Regional Centre of Excellence for High Performance Computing hosted at Loughborough University. We provide pay-as-you-go industrial access to a 3,008core Bull Linux cluster with high I/O, low-latency, unrestricted disk-storage, and free technical support.

[www.hpc-midlands.ac.uk](http://www.hpc-midlands.ac.uk)

Tel: 01509 223674

Email: [info@hpc-midlands.ac.uk](mailto:info@hpc-midlands.ac.uk)

@hpc\_midlands

/company/hpc-midlands

[google.com/+hpc-midlandsacuk](https://plus.google.com/+hpc-midlandsacuk)