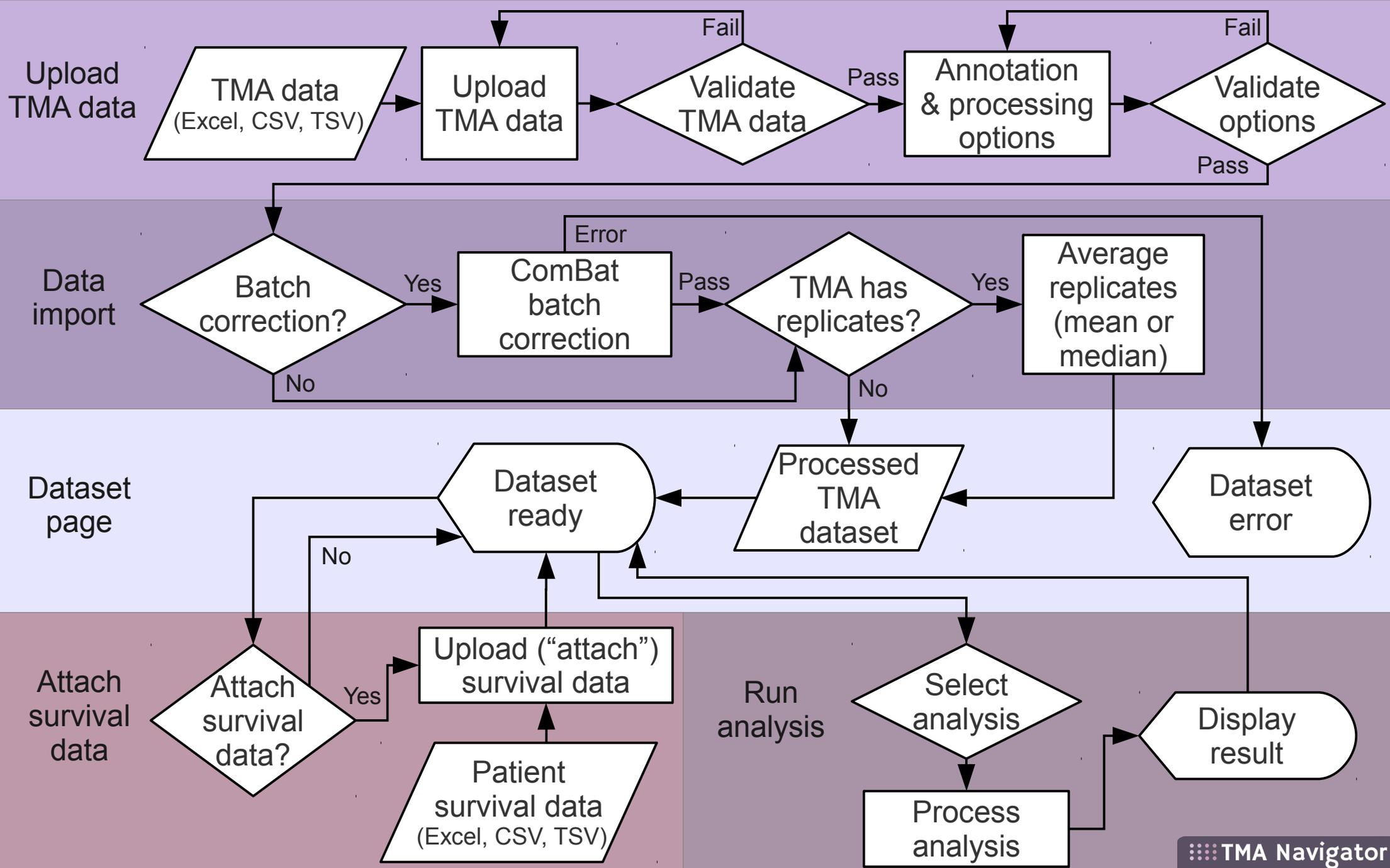


Supplementary information for “TMA Navigator: network inference, patient stratification and survival analysis with tissue microarray data”

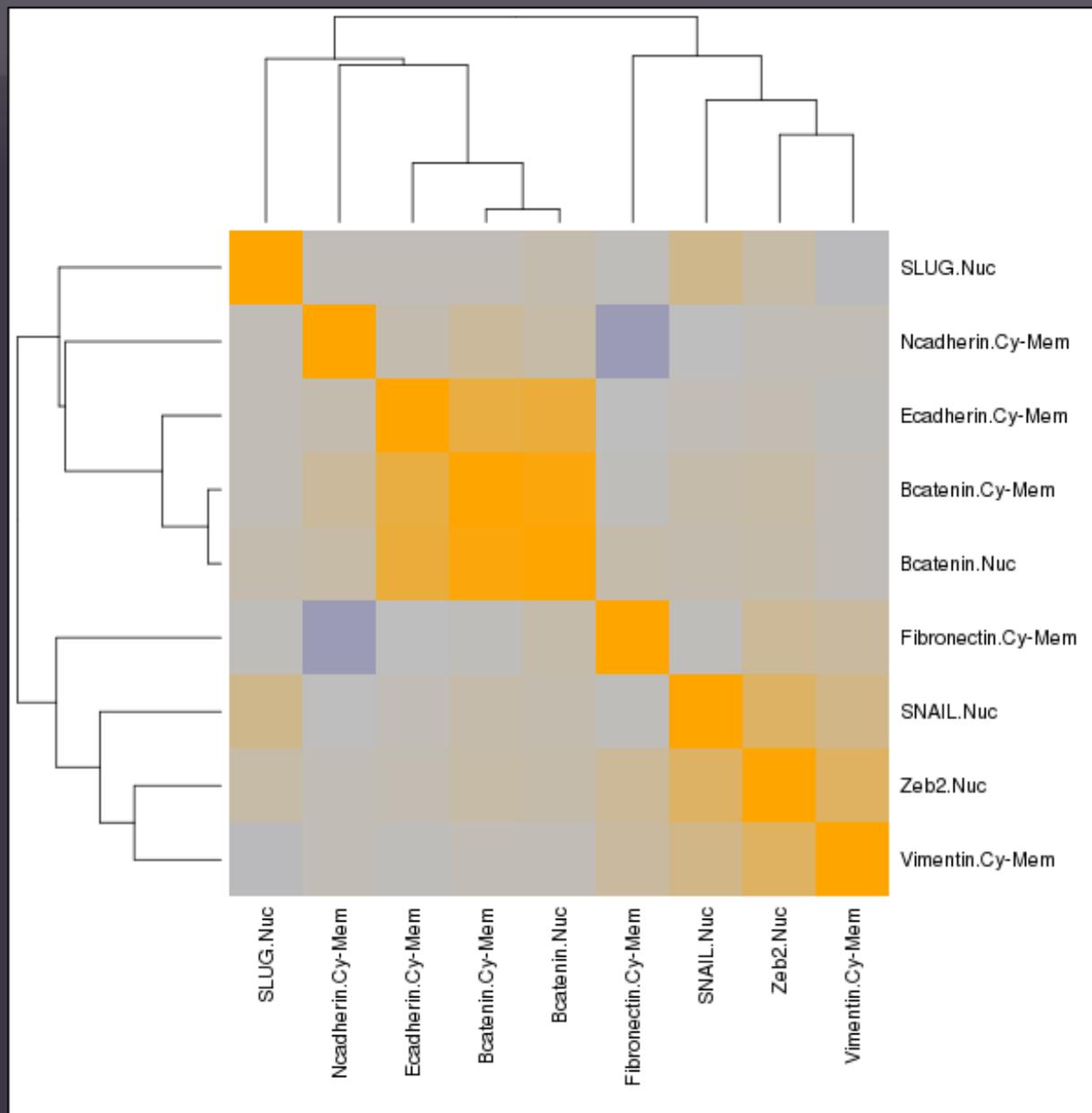
Alexander L. R. Lubbock, Elad Katz, David J. Harrison & Ian M. Overton

Contents

| | |
|-------------------------|---|
| Supplementary Figure S1 | TMA Navigator workflow |
| Supplementary Figure S2 | Marker heatmap |
| Supplementary Figure S3 | Stratification of trastuzumab treated breast cancers by tertiles of PTEN expression |
| Supplementary Figure S4 | TMA Navigator architecture |
| Supplementary Table S1 | List of antibodies used for AQUA immunofluorescence |



Supplementary Figure S1: TMA Navigator workflow. The figure has embedded hyperlinks which open to relevant sections of the user guide on the website. The first step in using TMA Navigator is to upload a table of marker scores. Following successful validation, data are annotated with a title and (optionally) notes. Options for score type (categorical or continuous), aggregation of replicates and batch correction are selected, as applicable. Data import with the selected options is initialised and the dataset added to the queue. Once processed, the dataset is ready for analysis from the 'dataset page'. At this stage, survival data may optionally be "attached" to the dataset by uploading a separate file. Analyses can be run from a dialog box accessed from the 'dataset page', and output is provided in a 'completed results' table. Datasets and results can be revisited using a unique key-protected URL (anonymous users) or by logging in (registered users).



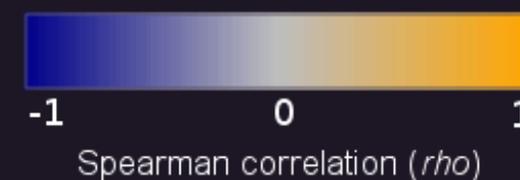
Heatmap

? Help

[Open in new window](#)

[Download as SVG image](#)

[Download correlations \(.tsv\)](#)



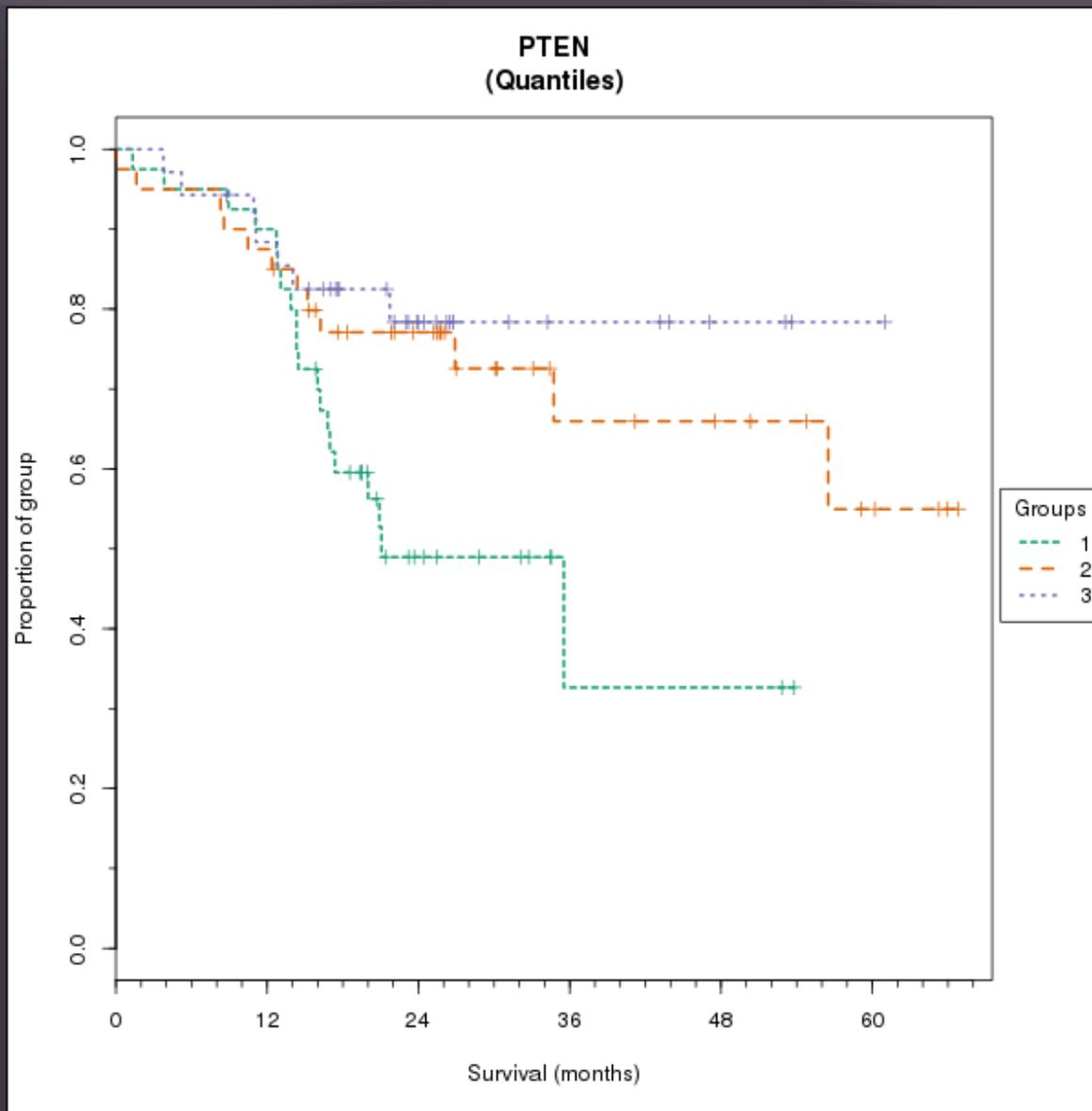
Supplementary Figure S2: Marker heatmap. Pairwise Spearman correlation is visualised as a heatmap (left) on a blue-grey-orange scale (right). The order of markers is determined by hierarchical clustering using complete linkage (far left and top-left). The main diagonal indicates self-correlation, always 1 by definition. The highest correlation cluster contains E-cadherin and β -catenin, which form an adhesion complex important in maintaining epithelial state. Snail, ZEB2 and Vimentin form another cluster; Snail and ZEB2 are EMT transcription factors which can both induce Vimentin expression, a known mesenchymal marker.

PTEN

pAKT.nuclei

pAKT.cyto

HER2



PTEN

[? Help](#)[Open in new window](#)[Download as SVG image](#)[Download groups \(.tsv\)](#)**Log-rank Chi-sq:**

7.817

FDR corrected p-value:

0.0207

Mean silhouette width:

0.5394

Group 1:

n: 40

Min: 0.00

Median: 103.33

Max: 133.33

Group 2:

n: 40

Min: 136.67

Median: 182.50

Max: 216.67

Group 3:

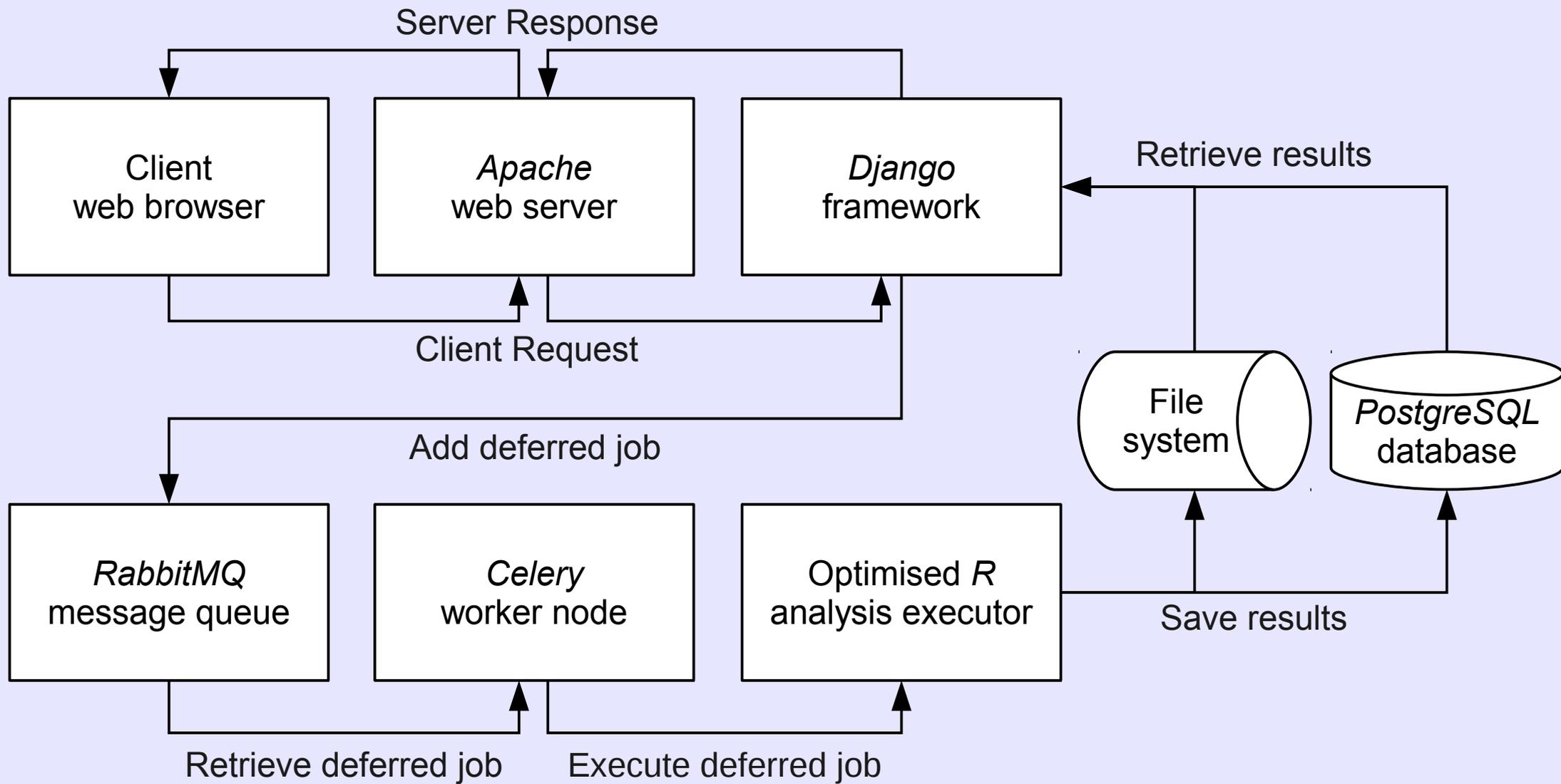
n: 35

Min: 220.00

Median: 273.33

Max: 300.00

Supplementary Figure S3: Stratification of trastuzumab treated breast cancers by tertiles of PTEN expression. Kaplan-Meier plot: x-axis denotes overall survival in months, y-axis the proportion of the group alive. PTEN has previously been shown to have a role in resistance to trastuzumab, and a significant positive relationship with survival was identified (FDR $p=0.0207$). The numbers of patients in each group is not identical due to the semi-continuous scoring system, where multiple patients take the same value at the group boundary.



Supplementary Figure S4: TMA Navigator architecture. The figure has embedded hyperlinks which open to the homepage of the relevant software, where applicable. TMA Navigator is a web application comprising several server-based software components. Users connect to the service using a web browser, which accesses web pages on an Apache web server. The site is written in Python using the Django web framework. Data are stored on the server file system and in a PostgreSQL database. Submitted 'jobs' – importing data, running analyses – are computed asynchronously using a RabbitMQ message queue system. When sufficient computing resources are available, a Celery worker node retrieves the next available job from the queue, if any, and executes it using a custom software library written in the R language. The results are saved to the file system and database for subsequent retrieval by the web server upon user request.

| Target | Source | Catalogue No. | Host |
|------------------|---------------------------|---------------|--------|
| E-cadherin | BD | 610181 | Mouse |
| Claudin7 | Abcam | ab75347 | Rabbit |
| N-cadherin | BD | 610921 | Mouse |
| Vimentin | Sigma | V 6630 | Mouse |
| Fibronectin | Abcam | ab2413 | Rabbit |
| Zeb2 | Acris | ZFHX1B | Rabbit |
| Zeb1 | Gift from D. Darling | n/a | Rabbit |
| Slug | LifeSpan Bio | LS-C30318 | Rabbit |
| Snail | Abcam | ab17732 | Rabbit |
| β -catenin | BD | 610153 | Mouse |
| MAL2 | Abcam | ab75347 | Rabbit |
| C35 | Vaccinex | (VX35) | Rabbit |
| Her2 | Dako | A0485 | Rabbit |
| ER-alpha | Vector | VP-E613 | Mouse |
| ALDH | BD | 611194 | Mouse |
| PTEN | Cell Signaling Technology | 9552 | Rabbit |
| pAKT | Cell Signaling Technology | 4060 | Rabbit |

Supplementary Table S1: List of antibodies used for AQUA immunofluorescence. Sodium citrate pH 6.0 was used for antigen retrieval.