# *Omics Central*, a data analytics dashboard: application to heart failure

**Amrit Singh**, Scott J Tebbutt, Bruce M McManus

https://amritsingh.ca

**PROOF Centre of Excellence, University of British Columbia, Vancouver, Canada**

## Introduction

- Preprocessing of high dimensional omics datasets differ but downstream analytical methods are often the same.
- Analytical methods such as those applied for exploratory data analysis, machine learning algorithms, gene set enrichment analysis, and resources for literature mining require coding knowledge in languages such as R, and python.

## Objectives

Develop an application:

- that connects commonly used tools in an omics analysis workflow (from biomarker discovery to literature searches) into one framework.
- that allows researchers to analyze omics datasets, share analyses and build reproducible workflows.

## Methods

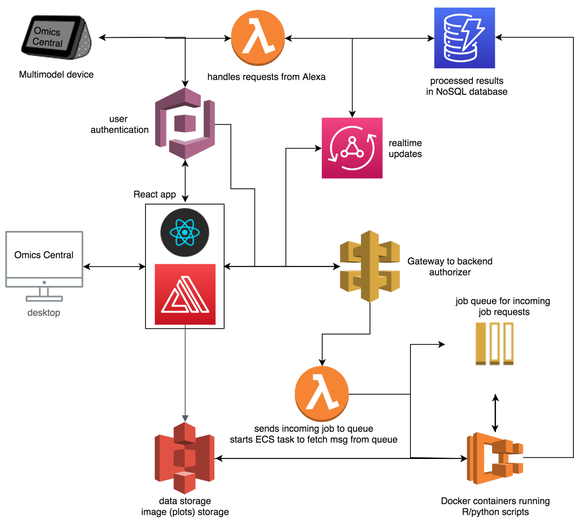


**Figure 1. Serverless application architecture using AWS.** The web app is a dashboard where users can upload their data and run various bioinformatics analyses. Users can also access various results using a custom Alexa Skill.

- This multimodal (web/voice) application (Fig. 1) is developed using serverless computing on Amazon Web Services (AWS):
  - Amazon S3 for hosting the website and data storage.
  - Amazon API Gateway as a REST API to make calls to the backend.
  - AWS Lambda to provide event driven operations.
  - Amazon Cognito user pool for user authentication and identity pool for user authorization.
  - Amazon DynamoDB as a NoSQL data store for processed results.
  - Elastic Container Service (triggered by Lambda) to run Docker containers consisting of R/python scripts for bioinformatics data analyses where the processed results (*e.g.* model parameters, classification performances) will be stored in DynamoDB.
- The front-end of the web application is developed using React.js with the Material design framework for styling.
- The voice application is built using the Alexa Skill Kit.
- The datasets used for the case studies have been previously published[1]:
  - Electrical variables (Holter monitors)
  - RNA expression (Affymetrix HU Gene 1.1 ST microarrays)
  - Protein expression (Multiple Reaction Monitoring (MRM) mass spectrometry)

## Results – Heart failure study

***Objective*: Predict 3-month hospitalizations**

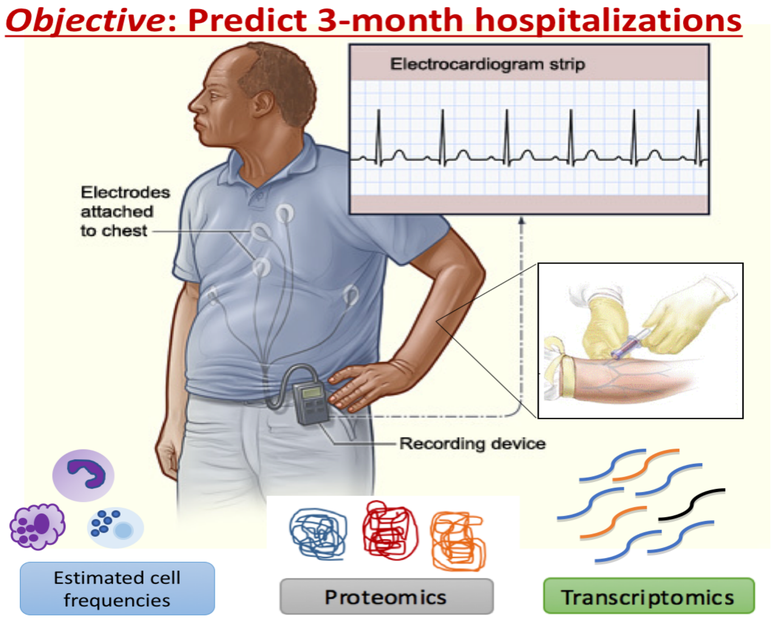


**Figure 2: Study schematic.** Baseline molecular data from the blood and electrical data from Holter monitors was used to identify predictive biomarkers of 3-month cardiac-related hospitalizations. Source: National Heart Lung and Blood Institute (NIH).

- 13 of the 58 patients had cardiac-related hospitalizations within 3-months of baseline blood collection and the first Holter monitor visit.
- Cell-type frequencies were estimated using statistical deconvolution (CIBERSORT[2]).
- The levels of BNP and creatinine were significantly elevated in baseline blood samples of hospitalized patients compared with patients who were not hospitalized.
- Differential expression analysis identified 1 cell type (monocytes), and 5 proteins ($\beta$-2 microglobulin [B2M], cystatin C [CST3], paraoxonase 3, apolipoprotein M, and apolipoprotein A2) with levels that were statistically different between patients who were hospitalized compared with those who were not at an FDR of 5%.

## Results – Interactive dashboard

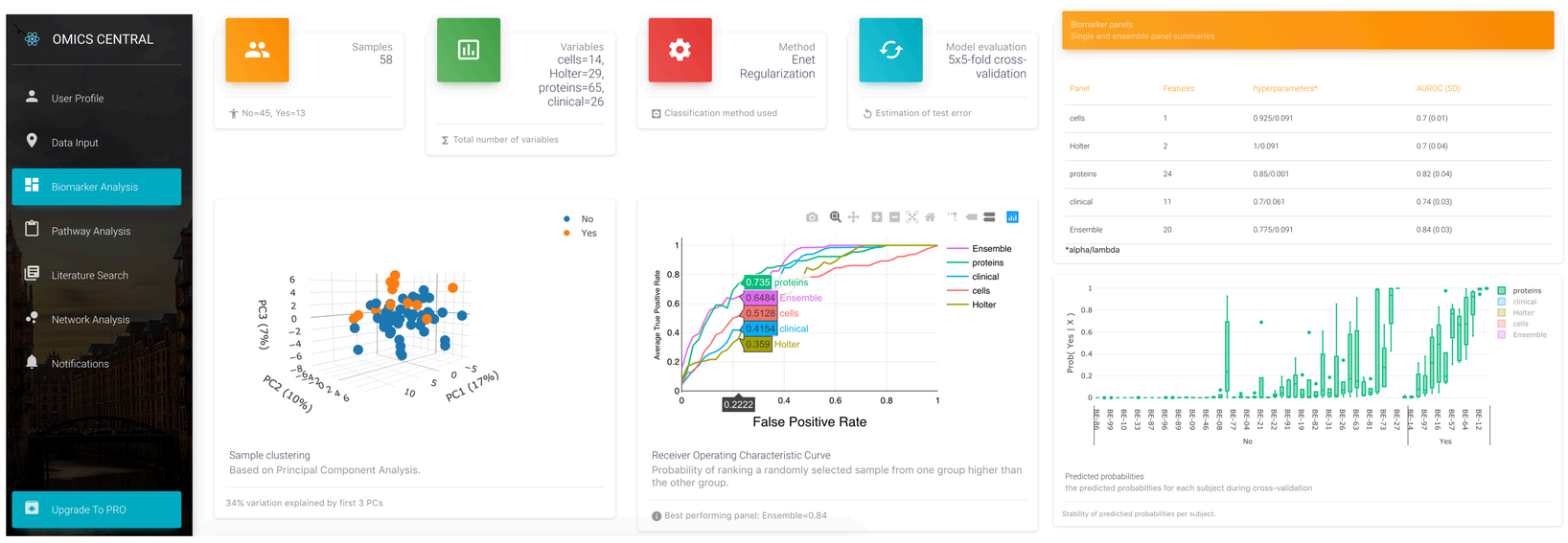


**Figure 3: Data analysis workflow.** Users can use this web-app on the go (mobile) or on a desktop browser. Current features include binary classification of one or more datasets (n samples x p variables) using Elastic Net regularization[3], as well as model evaluation using cross-validation. In the example above, 4 datasets (cell-type frequencies, Holter monitor data, protein expression and clinical variables) obtained on the same set of 58 patients were used to classify 13 hospitalized and 45 not hospitalized patients with heart failure. The Ensemble biomarker panel out-performed all other panels with an AUC=0.84.

## Discussion

- *Omics Central* is a graphical user interface that enables researchers to apply a pipeline of data analysis methods in order to identify biomarkers and implicate potential biological mechanisms in their condition of interest.
- In this study we use *Omics Central* to identify biomarker panels predictive of hospitalization status in patients with heart failure.
- The ratio between the daytime and nighttime Average heart rate was significantly higher in patients who were not hospitalized compared with patients who were hospitalized. The heart rate is usually lower during the night compared with during the day, however, in hospitalized patients the day- and night-time average heart rate were similar.
- Higher levels of B2M and CST3 have been associated with adverse cardiovascular outcomes in patients with coronary artery disease[4]. These different omics domains might be capturing different aspects of the HF etiology and therefore their combination resulted in improved classification performance compared with the usual clinical biomarkers.
- Future works include network analysis to incorporate curated relationships between biomarkers, and literature mining to generate additional hypotheses for downstream experiments.

## Limitations

- Small sample size, imbalance class sizes, lack of external cohort for model validation.
- The statistical deconvolution approach used to estimate cellular frequencies was originally validated using cancer samples, therefore it is difficult to determine its accuracy in our patient cohort.
- Very few patients had an ejection fraction ≥ 50%, therefore whether these results can be extrapolated to patients with HF with persevered ejection fraction remains a topic for further exploration.

## Funding sources

The authors thank Samsung and the Centre of Excellence for Prevention of Organ Failure for funding this research.

## References

1. Singh A *et al.*, **CJC** 2019;35(4):471-9.
2. Newman AM *et al.*, **Nat Methods** 2015;12:453-7.
3. Zou H *et al.*, **J R Stat Soc Series B Stat Methodol** 2005;67:301-20.
4. Nead KT *et al.*, **Am J Cardiol** 2013;111:851-6.