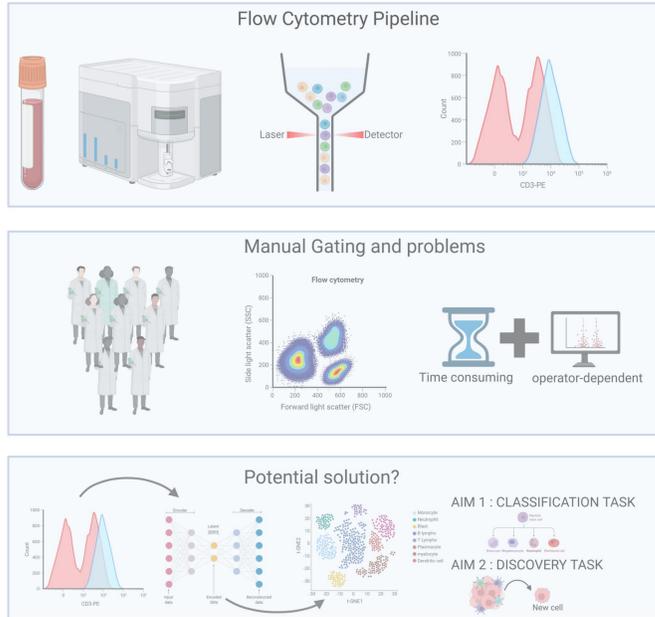


MARVIN: A Deep Generative Model for Flow Cytometry Analysis Informed by Biological Assumption

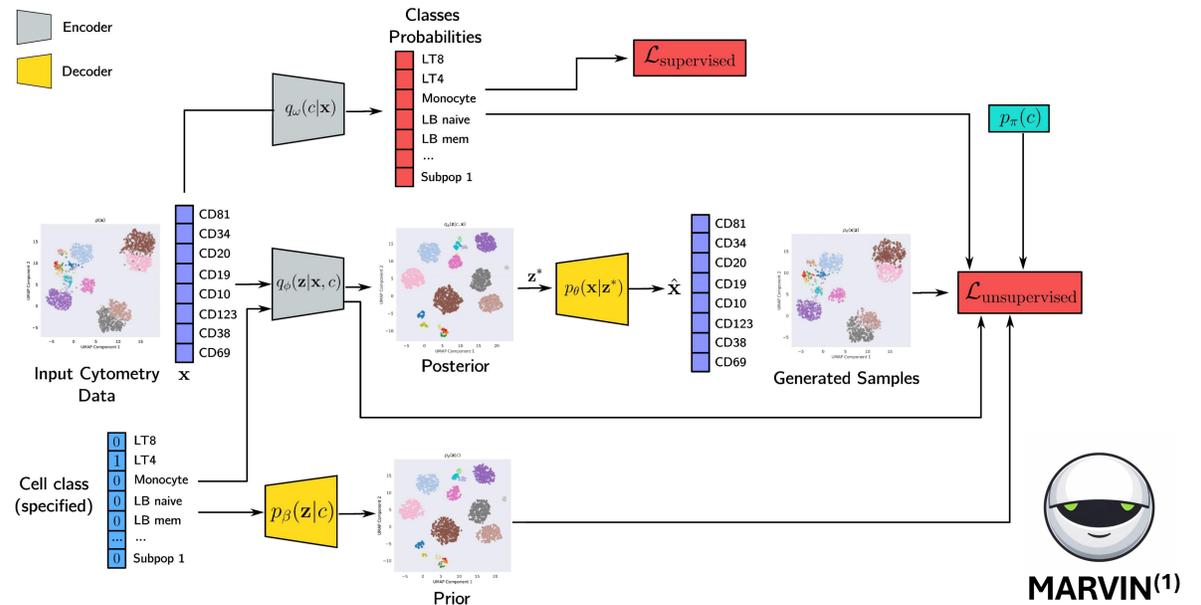
A. De Voeght^{1,2}, F. Bodart¹, F. Baron², G. Louppe¹.

- University of Liège, Montefiore Institute, Science with AI lab (SAIL), Liège, Belgium
- University of Liège, Centre Hospitalier Universitaire de Liège, Clinical Hematology, Liège, Belgium

Introduction & Objectives



Methods



Dataset. 5,480,065 cells (3 patients) and 10,222 tumoral cells (4 patients) with 12 annotated cell populations for a panel of 8 markers.

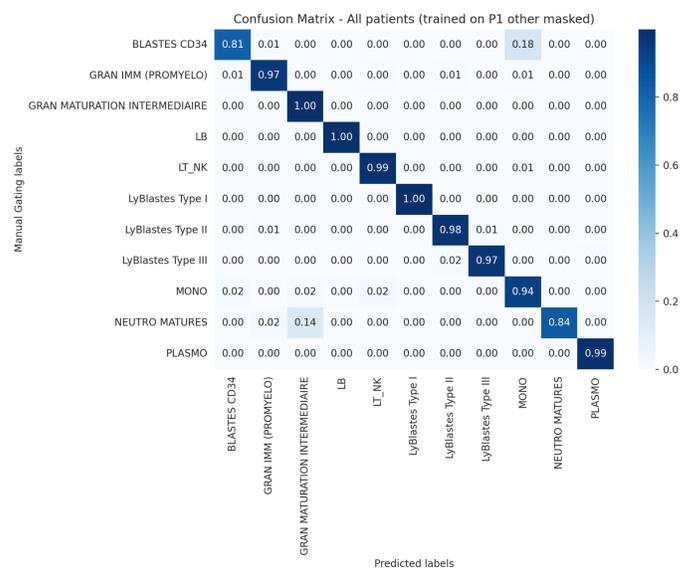
AIM 1 : classification task. Training on the first patient and applied on the second and third patient.

AIM 2a : subpopulation discovery. Evaluating whether MARVIN isolates pathological cells into additional clusters.

AIM 2b : anomaly detection. Measuring dissimilarity from known and unknown populations.

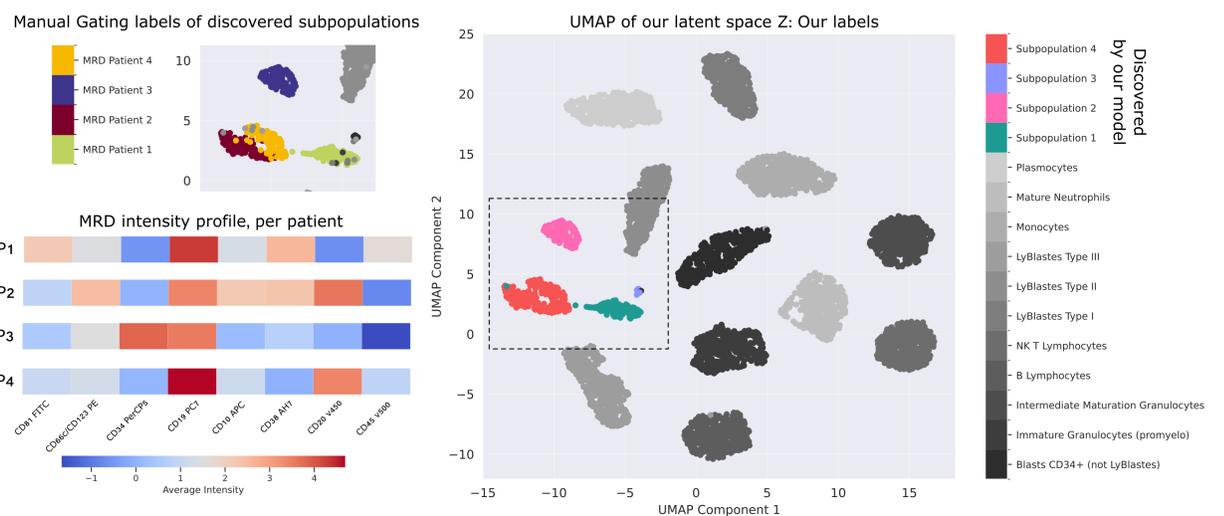
Results

Aim 1 : Classification task



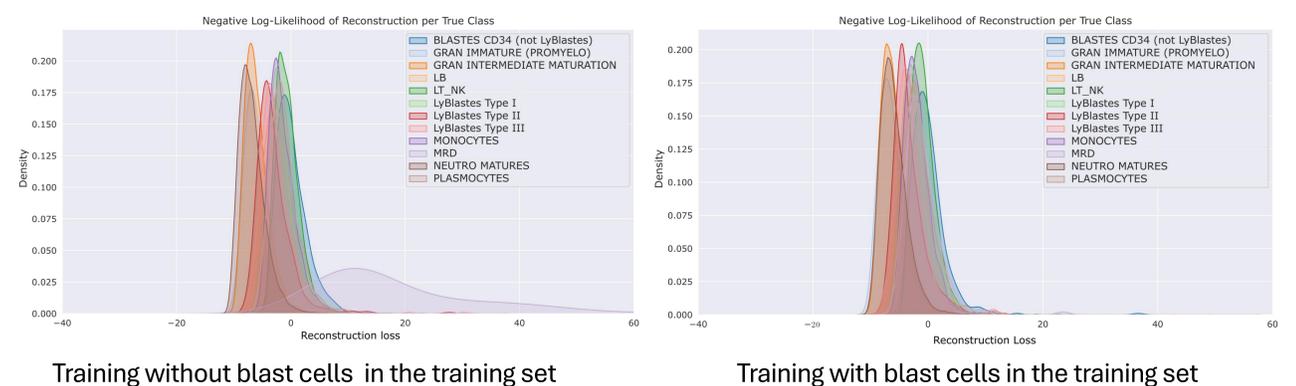
Aim 2 : Discovery task

Subpopulation discovery (during training)



MARVIN correctly assigned 99.2% leukemic cells in new clusters

Anomaly detection (after training)



Conclusions

FCM data, both standardised dataset and clinical dataset **can be automatically analysed**.

MARVIN is a **semi-supervised generative model** performing reconstruction and clustering, grounded in **biological assumption**.

MARVIN achieves **high classification accuracy** and **detects novel populations**.

Ongoing work focuses on biological refinement to improve rare population clustering and applying MARVIN to study MRD dynamics.