General Information

AutoFLOW aims at developing an **objective and automated tool** for multi-parameter FCM data analysis with robust and reliable MRD quantification

**Motivation**

**Labelled Data Set**
- 200 patients, treated according to AIEOP-BFM ALL 2009 protocol
- Gated in FACSDiva
- Hierarchical polygon filters
- Training set

**Representation**
Compact representation for data sets with millions of events needed
- reduce data
- keep relevant information

Methodology

**NMF based probabilistic model**
1) Relevant populations are modeled separately by GMM (no blast event is missed in the modeling phase)
2) Dimensionality reduction by NMF
3) Represent new input sample by linear combination of stored samples & assign each event of input sample to target populations using Bayes decision

Results (typical examples)

**MRD Predictions**

<table>
<thead>
<tr>
<th>Ground truth</th>
<th>NMF 20 factors</th>
<th>NMF 140 factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Events: 104306</td>
<td>Assigned as Blasts: 44.66%, PP: 931, FN: 283</td>
<td>Assigned as Blasts: 43.89%, PP: 528, FN: 701</td>
</tr>
<tr>
<td>Overall Events: 317556</td>
<td>Assigned as Blasts: 0.040%, PP: 36, FN: 11</td>
<td>Assigned as Blasts: 0.061%, PP: 54, FN: 10</td>
</tr>
</tbody>
</table>

Operator vs. AutoFLOW

**AutoFLOW MRD** n=102 unseen day 15 samples

Outliers (marked in red) occur due to:
- different phenotypes
- specific properties of the samples not represented by training data

Outlook
Model with additional explicit parameters

Conclusion

- Automated Flow Cytometric analysis of Acute Lymphoblastic Leukaemia (ALL)
- Ready-to-use FCM-based MRD-assessment tool for daily clinical practice for patients
- Reduce subjectivity, increase result comparability

Reference