**Announcement – FlowCAP-IV**

The FlowCAP Coordinating committee is pleased to announce the opening of the FlowCAP-IV competition.

The first three FlowCAP competitions (http://flowcap.flowsite.org/) consisted of a challenge phase, the distribution of standardized datasets and analysis by algorithm developers, followed by a conference at the NIH to discuss the results of the competition and plan future competitions.

The combined results of FlowCAP-I and FlowCAP-II were published (Nature Methods 10:228-38, 2013), and made the point that automated algorithms are rivaling manual analysis for some flow cytometry datasets. FlowCAP-III had three successful challenges, and the data from these are still being processed and extended, with the intention of publishing these results in paper(s) summarizing the outcomes of FlowCAP-III & -IV as part of a Special Issue of *Cytometry A* (see below).

**FlowCAP-IV: FlowCAP-in-CYTO**

For FlowCAP IV, we are changing the format in several ways. The results of the first FlowCAP meetings suggested that several algorithms are capable of improving on manual analysis, and so we believe that it is important to bring the algorithms to the attention of potential users (i.e., researchers with large and complex datasets that need objective and deep analysis). To achieve this objective, we are pleased to announce that we have arranged for FlowCAP sessions to be incorporated into the CYTO 2014 meeting in Fort Lauderdale, May 17-21, 2014.

The CYTO partnership offers many advantages. The CYTO meeting provides an opportunity to present the results of FlowCAP-IV to a much larger audience of potential users, it will allow algorithm developers to demonstrate selected analysis methods to potential users in a pre-meeting tutorial, and will allow us to learn from the other technical and biological content of the large CYTO meeting.

In addition to folding the FlowCAP meeting into CYTO, we have also made several additional changes. Only one challenge is being given - this will allow developers to focus their efforts. This challenge is related to one of the challenges in FlowCAP-III, but is based on a different dataset. The challenge is to predict the time to progression to AIDS among a cohort of HIV+ subjects, using antigen-stimulated peripheral blood mononuclear cell (PBMC) samples analyzed with a 14-color panel (see below).

In addition to the central challenge, we would also like to broaden the scope of FlowCAP to encourage presentations on innovative algorithmic approaches for the analysis of flow cytometry data in general (even if not used for the challenge) and new information arising from automated analysis. We would therefore encourage the flow analysis community to submit abstracts to CYTO, not only for poster presentations, but also for potential inclusion as oral presentations during the FlowCAP sessions (see structure below).

*Cytometry A* has expressed an interest in publishing a special issue of the journal focused on flow cytometry data analysis methods. Papers for this issue would be subject to the usual peer review process, but publishing together in a special focused issue will increase visibility of the papers. In addition to individual papers (for example, expanded from CYTO posters), the FlowCAP consortium will submit a paper summarizing the outcomes of the FlowCAP-IV

competition to the special issue, as well as the two papers we expect to arise from FlowCAP-III.

**FlowCAP-in-CYTO organization summary:**

1. One of the CYTO plenary sessions on data analysis will include a talk on flow cytometry analysis and the FlowCAP competition.
2. A parallel session will include presentations of different algorithmic approaches used for flow cytometry datasets, including the FlowCAP-IV challenge. Parallel session presentations are drawn from oral abstracts, ***so please submit abstracts by January 22 if at all possible.***
3. A three-hour workshop will discuss the results of the FlowCAP competition, and plan for future FlowCAP competitions.
4. A tutorial/demonstration session, the day before the main CYTO meeting, will demonstrate how to use some of the algorithms.
5. A special issue of *Cytometry A*, will bring together papers from FlowCAP-III and IV. Current plans are that papers would be solicited after the CYTO2014 meeting, submitted in the Fall, reviewed and then published in Spring, 2015.
6. The current CYTO deadline for late-breaking abstracts is March 11. This is for posters only, but we would like to encourage posters on your solutions to the FlowCAP IV challenge, to provide additional discussion.

## FlowCAP-IV:

**Introduction**

Below are instructions for participating in the FlowCAP-IV challenges. For more information please register at http://flowcap.flowsite.org/, join the Google Group and mailing list at http://groups.google.com/group/flowcap, or contact rbrinkman@bccrc.ca or Tim_Mosmann@urmc.rochester.edu. Please forward this announcement to any colleagues who might be interested in participating.

**Data Usage Agreement**

Publishing or using the datasets provided by FlowCAP for other purposes is prohibited until the FlowCAP Consortium publishes the results. At that point, the datasets and results of the FlowCAP challenge will be made publicly available for any use after publication. Once submitted to FlowCAP, the software and results cannot be withdrawn, but can be de-identified upon submitter's request.

**The Challenge:**

The goal of this challenge is to identify features in high-dimensional flow cytometry data (e.g. cell populations) that most closely correlate with a clinical outcome (in this case, progression to AIDS in HIV+ subjects) and can be used to predict the outcome in test subjects. The dataset includes PBMC samples that have been stimulated with HIV antigens, and their unstimulated controls. For each sample, data for FSC-A and FSC-H (for doublet removal), V-Amine/CD14 and SSC (for gating live lymphocytes), and 12 other markers are provided. It is known that there are predictive populations, but these are subtle and so this is expected to be a difficult challenge. On registering for the FlowCAP competition at http://flowcap.flowsite.org/, you will be provided with a link to the dataset and meta-data.

For each HIV-infected patient data from an unstimulated and an HIV antigen-stimulated sample is provided. The first meta-data spreadsheet (FlowCAPMetaData.csv) links:

A: Status - the observed clinical status (1 = progression to AIDS or death, 0 = no progression or death – NOTE – corrected 2/27/2014);
B: Progression/Survival Time - either the time to onset of AIDS for each subject (in days) or the time to last evaluation (NOTE: a few samples have negative values, i.e. the sample was taken shortly after diagnosis of AIDS);
C: Stim – FCS file names of samples stimulated with antigen: and
D: Unstim – FCS file names of unstimulated samples.
Half of the labels are provided for training purposes.

The second meta-data file (FlowCAPchannels.csv) contains a key indicating the antibody specificities of the 13 fluorescence channels used – the remaining 5 fluorescence channels were not used. These parameters were uniform across all samples although the labeling is not always consistent between files.

**Information required for each submission** (submit to flowcap@flowsite.org by March 11, 2014):
1. Your main task is to predict the "Progression/Survival Times" (i.e. AIDS-free survival). Please include the times (in days) predicted by your model for both the training and test sets, i.e. all 766 files. The values in this column can be either the predicted number of survival days or a correlate of it (e.g., values between 0 and 1 with 1 corresponding to the highest survival time). Your results will be primarily evaluated using a Cox proportional hazards regression and a log-rank test on the test set (the samples with missing survival information). However, we do ask for prediction for the samples in training set as well.

2. Please provide the 'feature matrix' of the values used for your predictive analysis. This needs to be a comma separated spreadsheet with features as columns and patients as rows (as ordered in the meta-data spreadsheet).

3. If possible, please also submit a description of the most predictive cell populations or features as identified by your analysis. This information may have a different format depending on the methods used by each algorithm, but for example could include the marker expression (positive/negative or high/medium/low) profiles for informative populations used to make the predictions; or medians (in original units) and cell numbers of the most relevant sub-populations; or any other features used to arrive at the results. This is an optional item.

4. Please also describe the algorithmic approach in general terms (see Nature Methods 10:228-38, 2013 for examples), and indicate the availability of the programs.

5. Reproducibility: Well-documented source code that can be used to reproduce your results from the raw data is mandatory. Please note that FlowCAP no longer accepts pseudocode or binary files. Please provide detailed instructions for proper configuration of your scripts (paths, libraries for parallel processing, internal variable values, etc.). The FlowCAP evaluation committee reserves the right to re-run any analysis to reproduce submission results.

**Timeline**

| | |
|---|---|
| January 20, 2014: | Call for participants. |
| March 11, 2014: | Deadline for submitting results: |
| May 17-21, 2014: | FlowCAP-IV Summit to be held during the CYTO 2014 meeting, Fort Lauderdale, Florida USA. |