Stockholm University

Mathematical Statistics
Stockholm University
Master Thesis **2012:6**
http://www.math.su.se

Quality control and analysis of flow cytometry data

Amos Thairu*

June 2012

Abstract

Data captured from new generation flow cytometers are characterised by increasing volume and complexity. For example a single specimen may give rise to a data set of about 1 million rows. However, analysis of these data is usually performed using user-interactive software which makes it time consuming and highly affected by operator experience. The main tasks involved in flow cytometry data analysis include visualization of scatter plots and extraction and summarization of cell sub-populations. In an attempt to solve the problems associated with the user-interactive methods, various software developments have recently emerged. In this project statistical methods are implemented in R and Bioconductor to automate the analysis of flow cytometry data collected from a HIV research study. The results are compared with those obtained by an expert analysing the same data using the traditional methods. The R functions developed are applied to an analysis of specimens taken from a HIV-infected infant at birth and at 3 months of age and changes in cell populations are described. The R code developed would be useful for laboratory scientists for automating some of the steps in the analysis of flow cytometry data, thereby offering significant savings in time and more reproducible results.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: mbuguatea@yahoo.com. Supervisor: Niklas Norén.