# Compositional challenges in the analysis of flow cytometry cell count data.

**J. D. Silverman**[1], **J. Smith**[2]
[1]Duke University, NC, USA; *justin.silverman@duke.edu*
[2]Duke University, NC, USA; *js.smith@duke.edu*

**Abstract**

Flow cytometry is a widely utilized method of cell counting, cell sorting, and quantifying the relative abundance of different cell populations. The chemical and biophysical characteristics of cells are quantified by both inherent properties as well as by fluorescent labeling of cellular components. Common Populations of different cell types can be quantified by grouping cells based on these measured properties and fluorescent spectra. Thus, each sample is ultimately represented by a vector of counts of different cell types. Importantly, the total number of cells measured for a given sample is often fixed or arbitrary (unrelated to a meaningful biological quantity); this feature complicates the analysis of flow cytometry count data by introducing compositional constraints. To demonstrate the potential artifacts that can arise in analysis of this type of data, we investigate the relative abundance of three cell types (T-cells, Neutrophils, and Keratinocytes) within a mouse ear either at baseline or after exposure to a dermal allergen. In response to this dermal allergen, we expect the lymphocytic cells (T-cells and Neutrophils) to increase in abundance, whereas the Keratinocytes, having little immunologic role, should remain constant. We demonstrate that naive analyses based on either the raw counts measured for each sample or the inferred proportions can lead to spurious results (e.g., a decrease in the level of Keratinocytes in response to a dermal allergen). In contrast, we demonstrate how the use of simple tools from the field of compositional data analysis can avoid these spurious conclusions and result in the correct interpretation of the data.