**Number of Events to Collect for Flow or Mass Cytometry**

**Adapted from:** <https://utahflowcytometry.wordpress.com/2013/07/26/how-many-cells-should-i-count/>

[Technical issues: flow cytometry and rare event analysis.](https://pubmed.ncbi.nlm.nih.gov/23590661/)

Hedley BD, Keeney M.Int J Lab Hematol. 2013 Jun;35(3):344-50. doi: 10.1111/ijlh.12068.PMID: 23590661

***We are often asked:***

How many events/sample should I collect in my flow or mass cytometry experiments?

***Here is the answer!***

The total number of cells you should collect is determined by:

1. **the frequency of the rarest population** you want to measure ***and***
2. **the precision**, represented as the coefficient of variation (CV), with which you want to measure it.

* The precision of your measurement is calculated using the Poisson distribution based on the number of events you count in your subset of interest.
* The essential feature of Poisson distributions is that if N events are observed in your target subset, then the standard deviation (SD) associated with that count is square root of N.
* The coefficient of variation (CV) is then given by: CV(%) = 100/sqrt N.
* Thus, as you count more cells in your target subset, the CV will decrease, indicating higher precision.

This calculation has **two important consequences** for determining how many events to collect:

1. If you want to detect a 1% population with the same precision (CV) as a 10% population, you need to collect 10 times more total events. In mathematical terms, if the subset measured is reduced by a factor of S, then you must process S times more total events to maintain the same precision.
2. To decrease your CV (and thus improve precision) from 5% to 2.5%, you will need to count 4 times as many cells in your target subset. In other words, counting 4X more total cells for your subset of interest will increase the precision of your measurement by a factor of 2.

The following a reference chart provides estimates for how many events to process to measure subsets of 10%, 1% or 0.01% with a range of precisions.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***For a CV of (%)*** | ***2.5*** | ***5*** | ***10*** | ***20*** |
| ***Events Counted (N)*** | ***1,600*** | ***400*** | ***100*** | ***25*** |
| ***Subset Frequency (%)*** |  | | | |
| **10** | 16,000 | 4,000 | 1,000 | 250 |
| **1** | ***160,000*** | ***40,000*** | 10,000 | 2,500 |
| **0.1** | 1,600,000 | ***400,000*** | ***100,000*** | 25,000 |
| **0.01** | 16,000,000 | 4,000,000 | 1,000,000 | ***250,000*** |

**For most applications CVs in the 2-5% range are considered ideal:**

* For a CV of 5%, you need to count 400 cells of your population of interest.
* For a 1% population, you will need to count 40,000 total events to achieve a CV of 5%.
* For a 0.1% population, you will need to count 400,000 total events to achieve a CV of 5%.
* Note that if your sample has many dead cells and/or doublets you need to consider the **% of total cells** collected represented by your target, **not the % of live singlets**.
* Lower precision (10-20%) is generally acceptable for rare event analysis when it is not feasible to collect enough events to achieve higher precision