FACS Advance Training Course

Concept of *Parameter Setting* and *Data Analysis*

尚博生物科技有限公司
www.cell-bio.com.tw
Glenn Yang
apoptosistw@gmail.com
02-27855860-0953062485

Outline

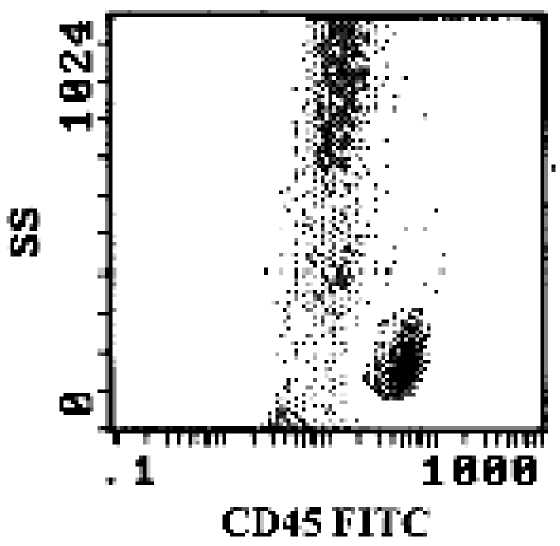
- *Region and Gate*
- *Statistic*
- *Positive vs. Negative*
- *Data Analysis (WinMDI 2.8)*
- *Professional Reporter (FCSExpress 3.0)*

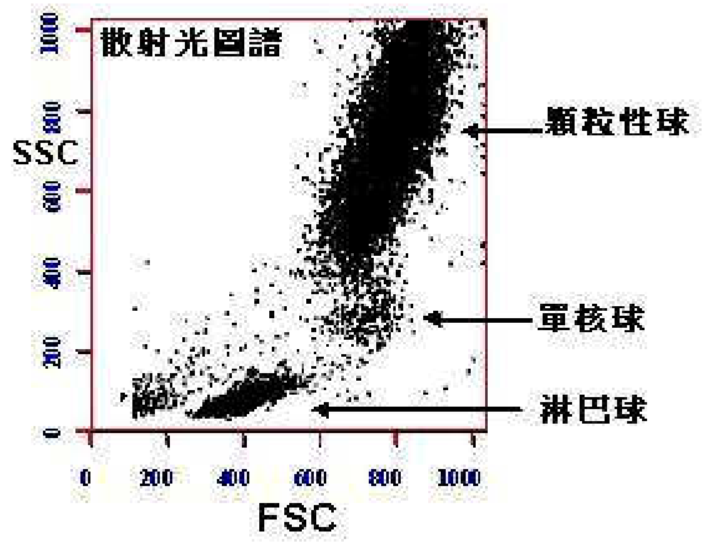
Cell-Bio

Why Region and Gate?

Particle (cell) Discrimination

- Problem :
  - Very often, samples are heterogeneous
    →there are events which are not of interest
    (other cells, debris, electronic noise).
  - Several clusters of interest mixed together

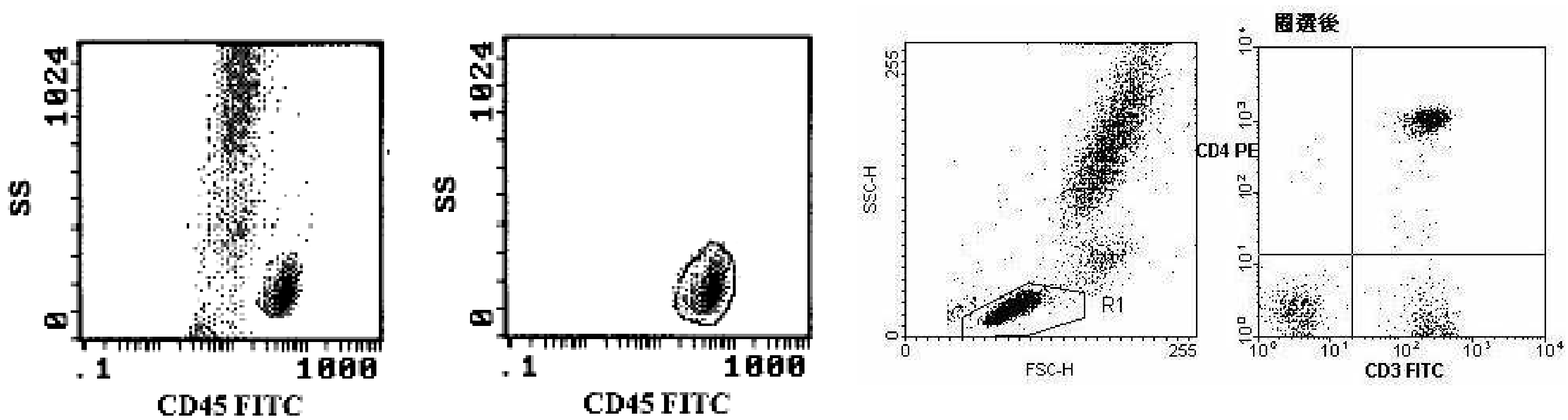




Why Region and Gate?

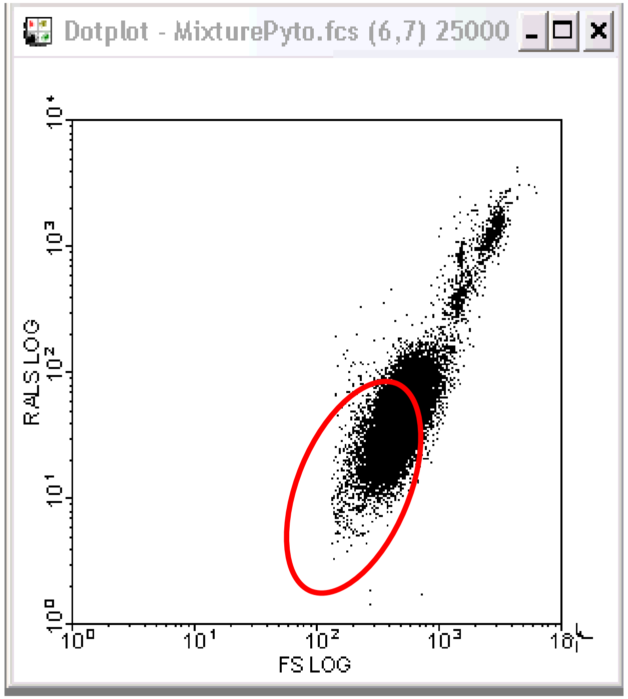
Particle (cell) Discrimination

- Solution :
  - Discriminate the cells of interest.
  - Need to exclude the unwanted events from the analysis.



Cell-Bio

What is a Region?

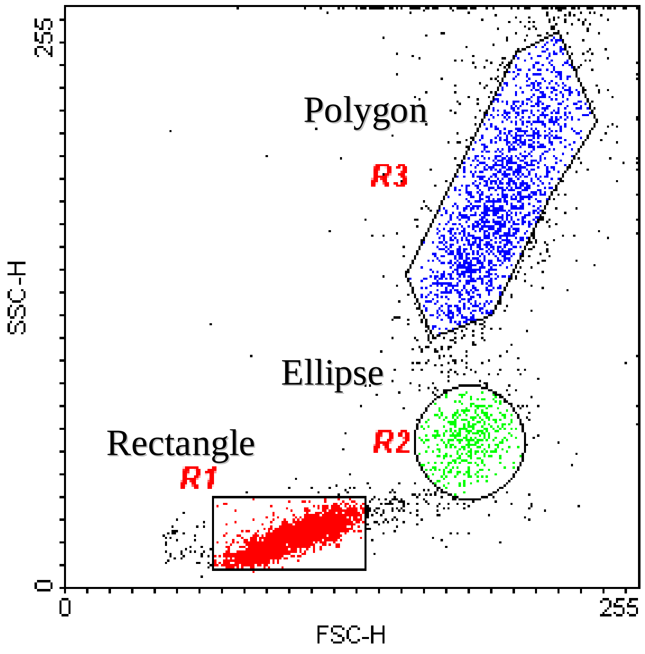


A **region** can be defined as set of points carefully selected by the user that determine an area on a graph.

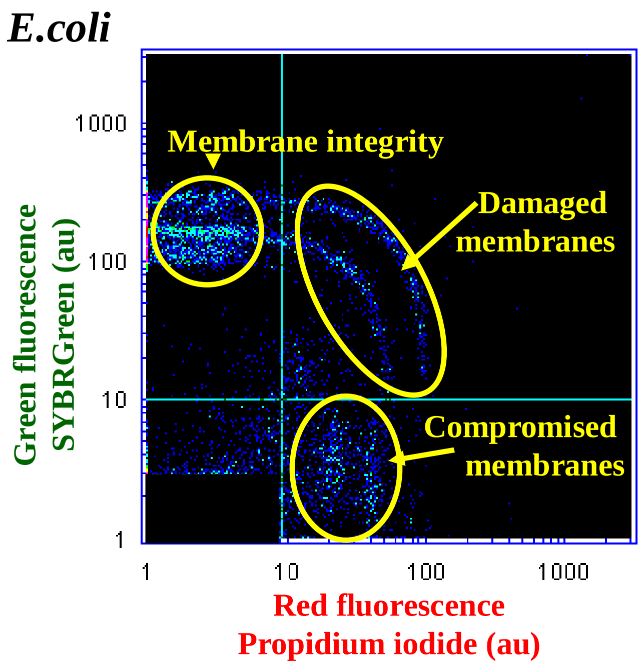
Several regions can be defined on the same graph.

➔ Isolate the cluster(s) of interest
➔ Better discrimination of the cluster(s) using color

Different styles of regions

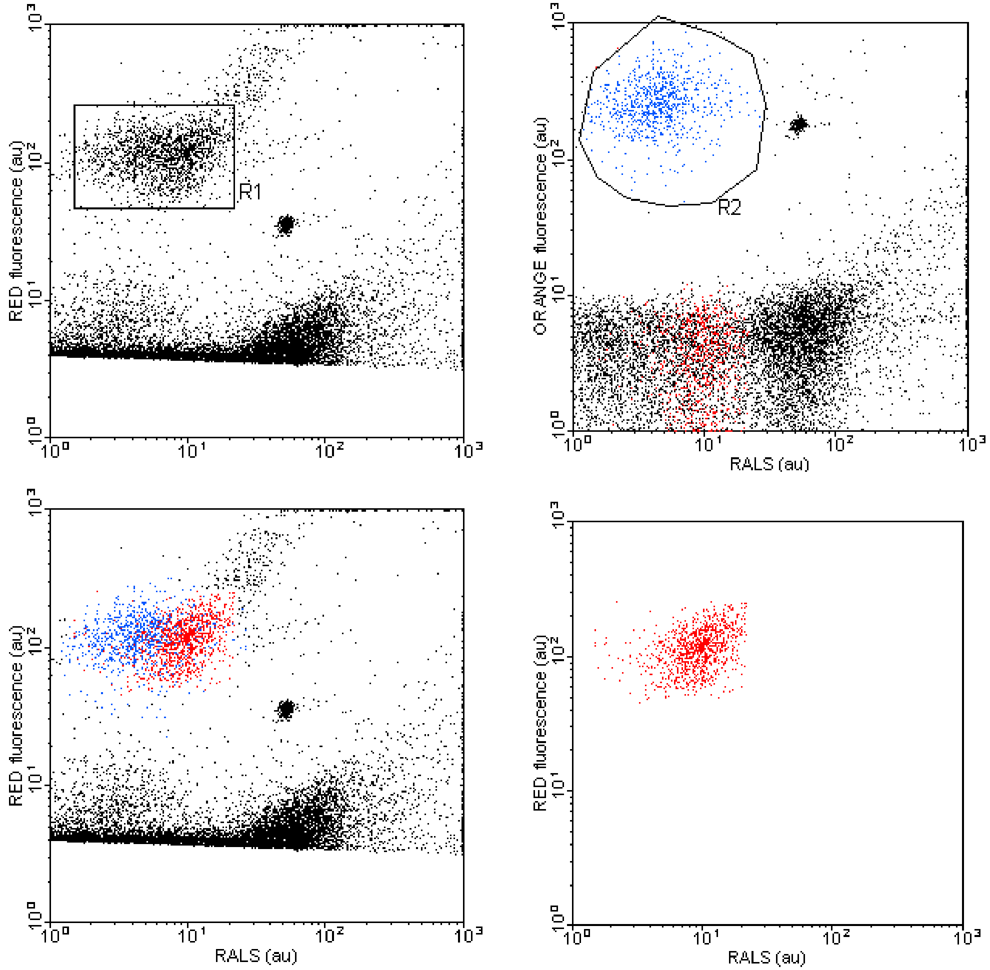


Cluster discrimination



Positive/Negative cell identification

What is a Gate?



A **gate** can be defined as one or more regions combined using Boolean (logic) operators (AND, NOT, OR)

Defines a subset of the data to be displayed.

- Used to compute **statistics** and characterize the subset of events selected

- Get rid of noise and save space on disks

Gating

- Real-time gating vs. software gating
- Establishing regions
- Gating strategies
- Quadrant analysis
- Complex or Boolean gates
- Back gating

Cell-Bio

Real-Time vs. Software Gating

Real-time or live gating:

-restrict the data that will be accepted by a computer (some characteristic must be met before data is stored)

Software or analysis gating:

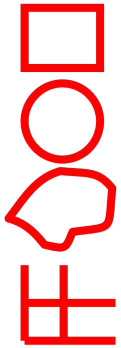
-excludes certain stored data from a particular analysis procedure

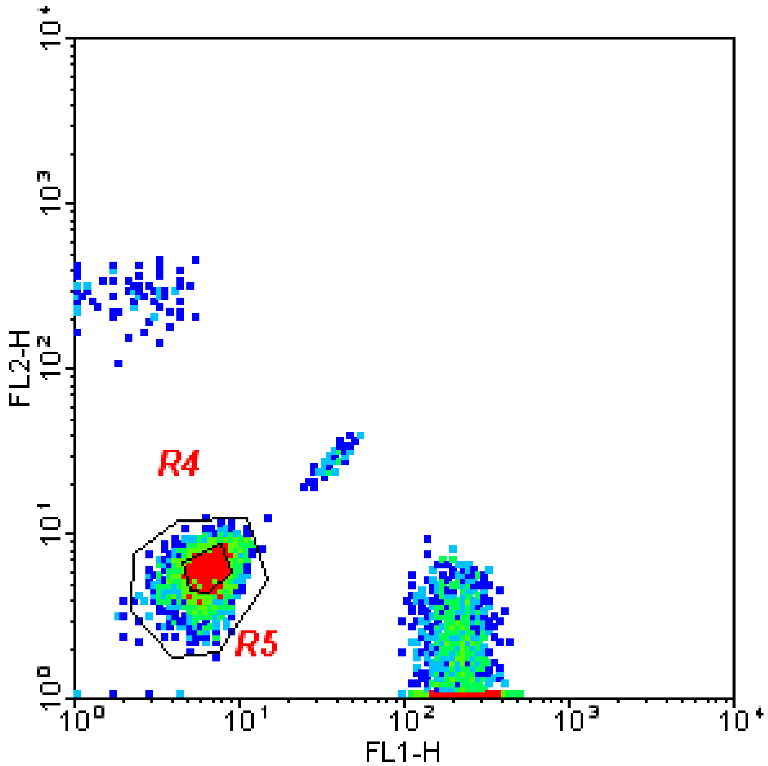
Cell-Bio

Establishing Regions

- Establishing regions:
  - -objective or subjective?
  - -training/skill/practice

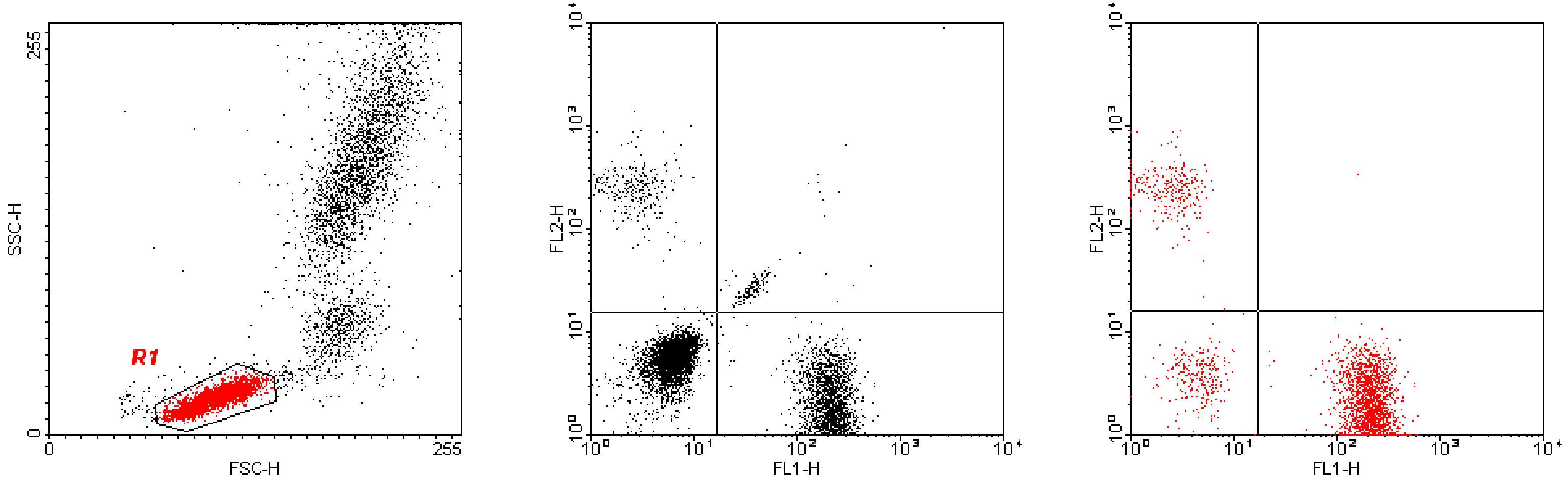
- Possible shapes:
  - -rectangles
  - -ellipses
  - -free-hand
  - -quadrants





- Statistics

Using Gates



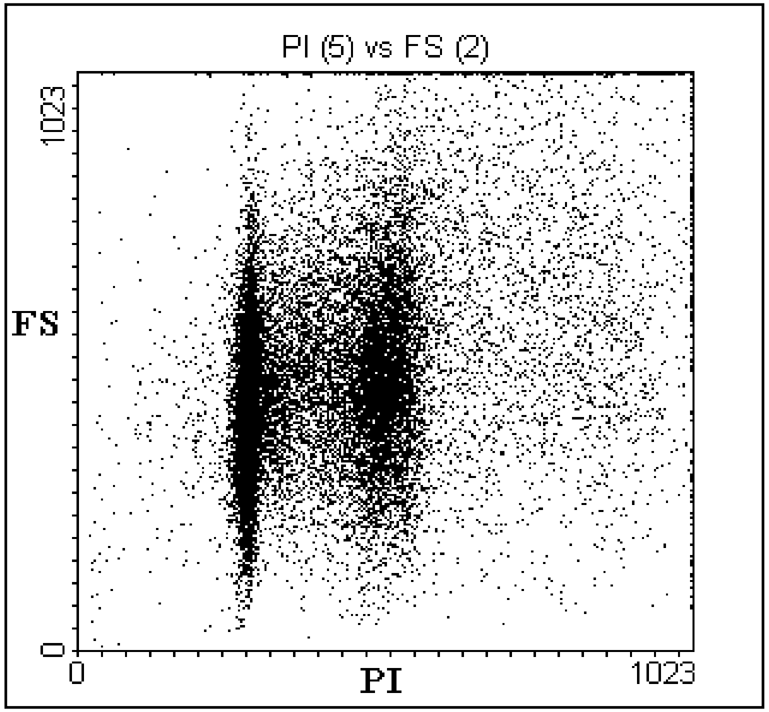
Region 1 established

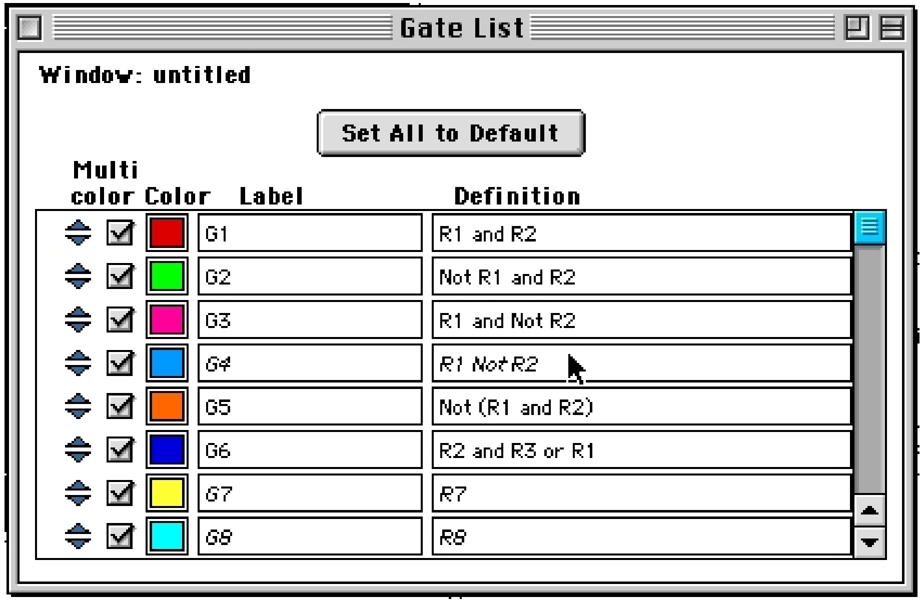
No Gated

Gated on Region 1

Complex or Boolean Gating

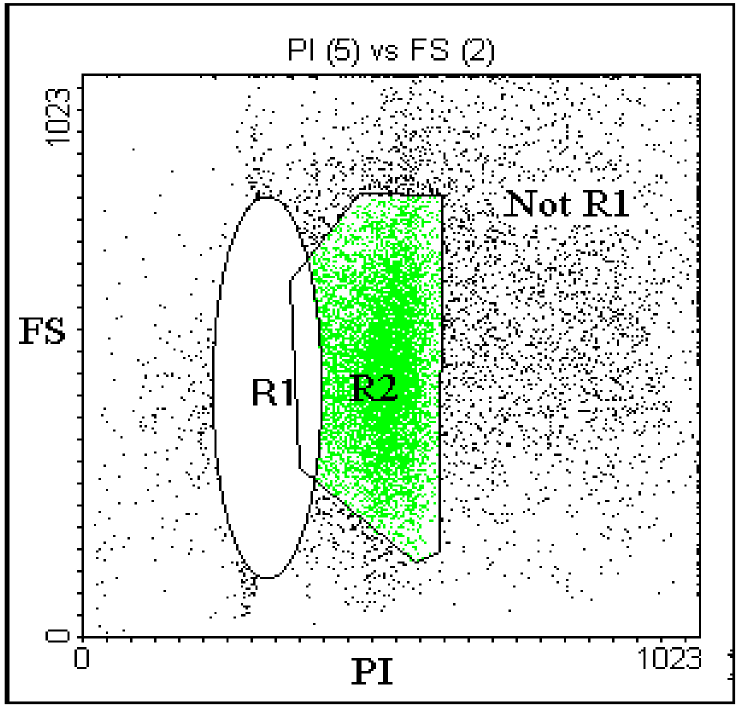
With two overlapping regions, several options are available:





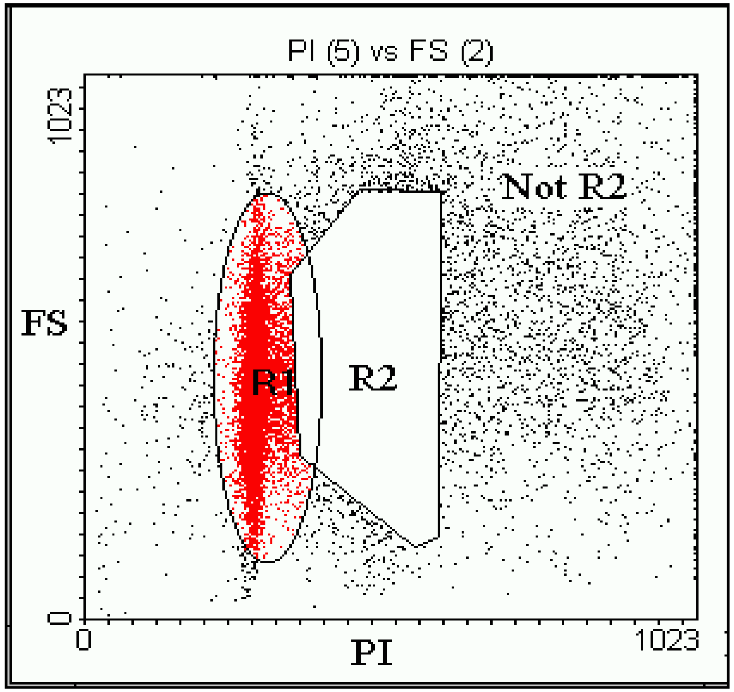
Boolean Gating

Not Region 1:



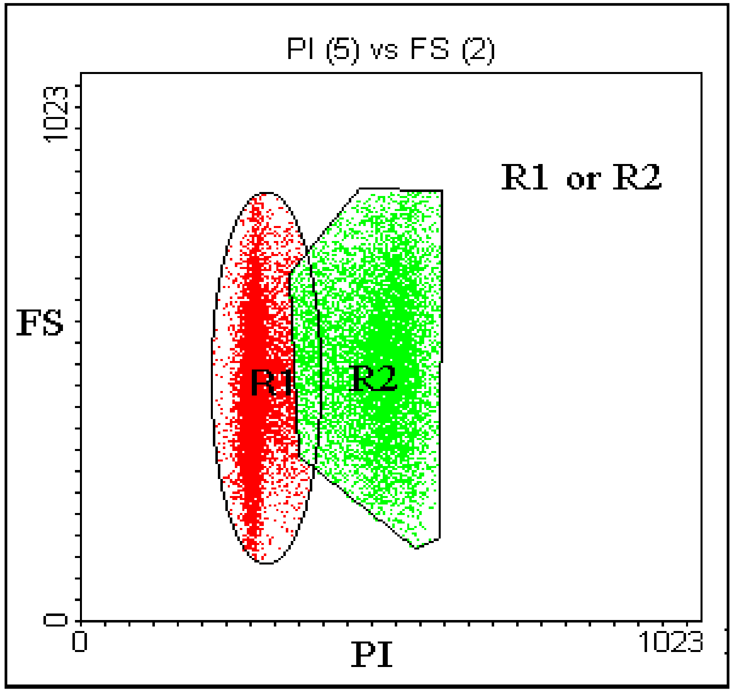
Boolean Gating

Not Region 2:



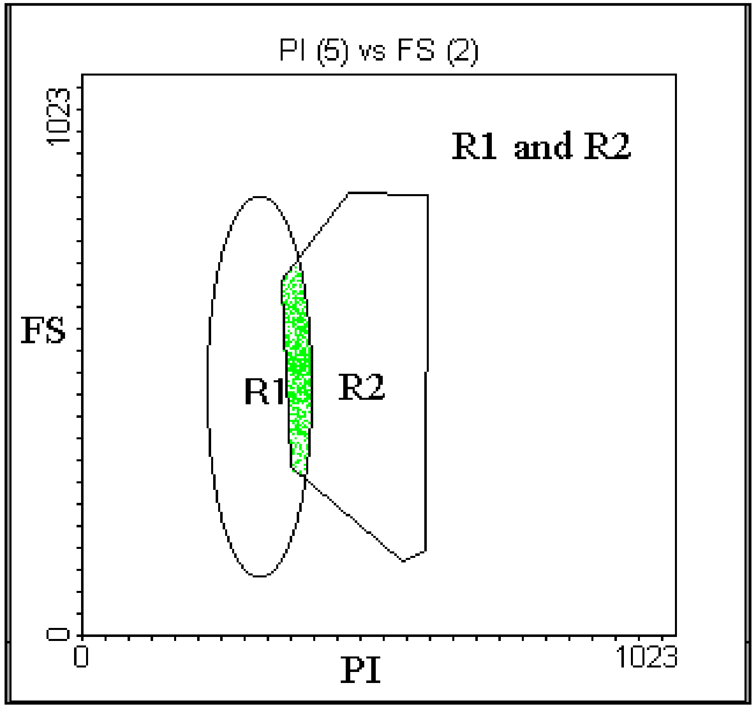
Boolean Gating

Region 1 or Region 2:



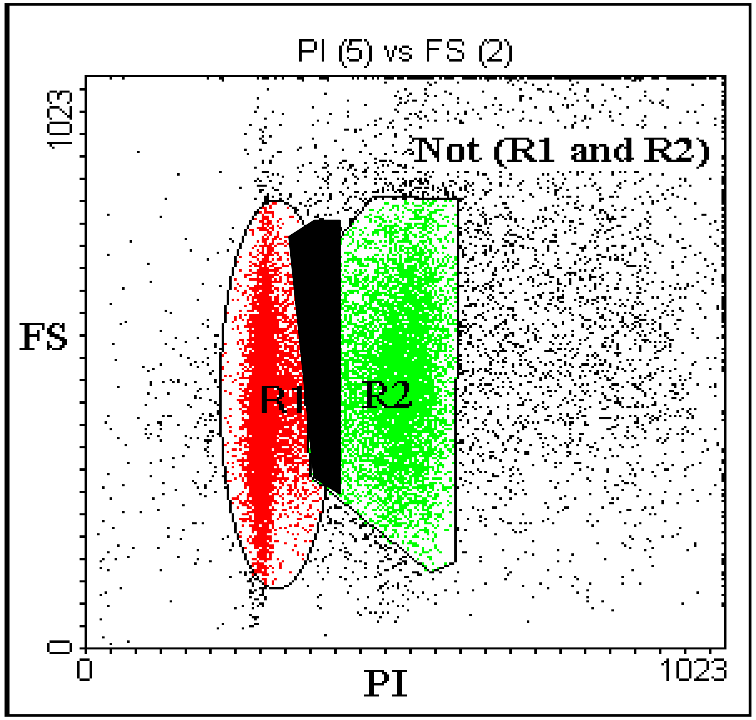
Boolean Gating

Region 1 and Region 2:

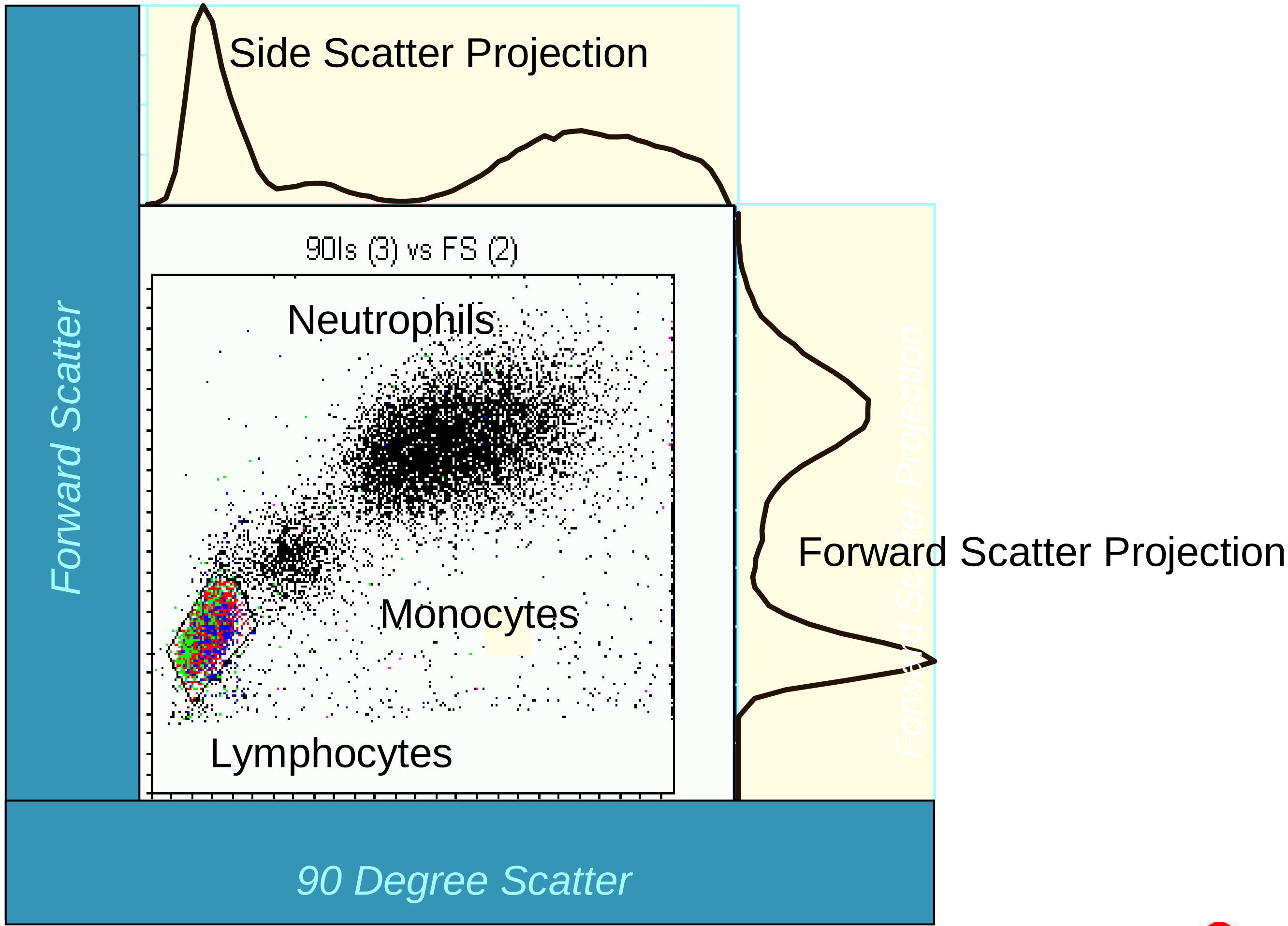


Boolean Gating

Not (Region1 and Region 2):

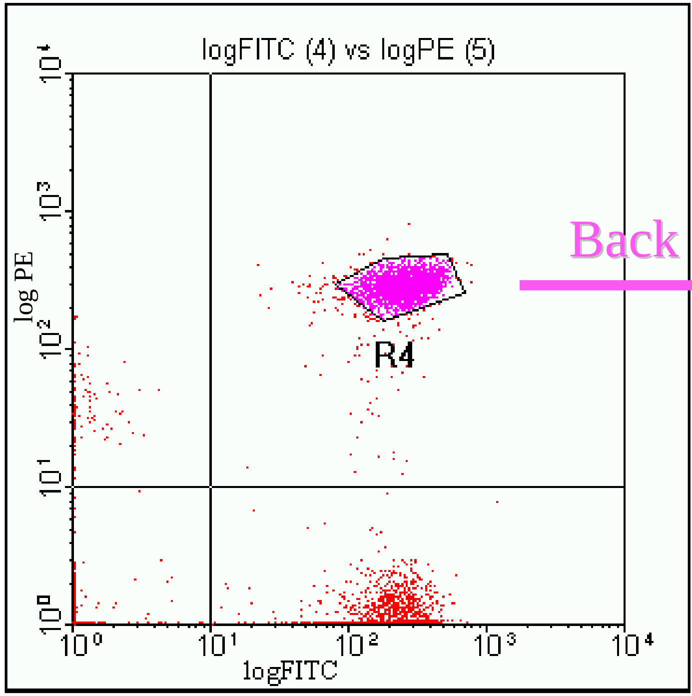


Light Scatter Gating

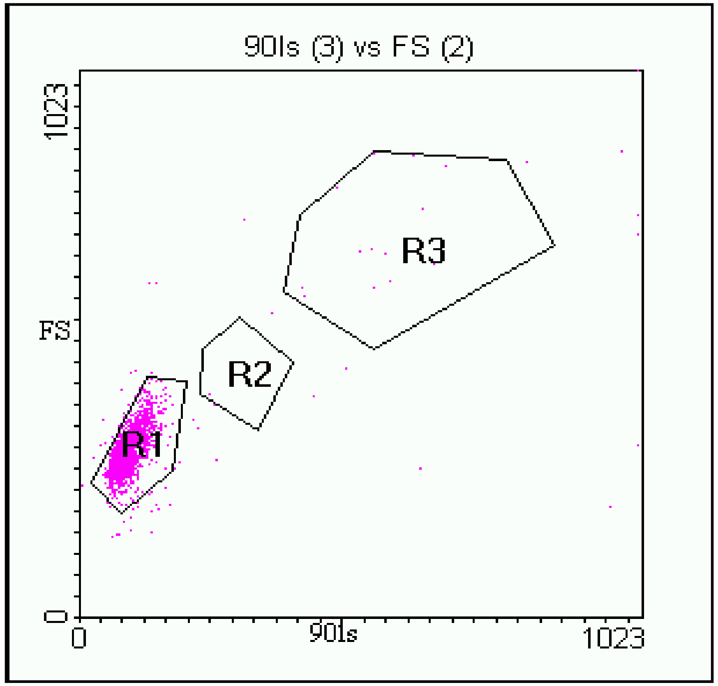


Human white blood cells

Back-Gating



Back gate



Region 4 established

Back-gating using Region 4

Statistics

Prior the statistical analysis of the clusters, consider these two factors :

1. **Sample size**:
The precision of the statistical analysis depends on the number of cells analyzed (Poisson Law → Std Deviation = $\sqrt{}$(n))
When the number of events increases the coefficient of variation of the estimate decreases.

2. **Incorrect choice of statistics** impacts the relevance of the results.

The mean(s)

The **mean** = one of the most widely used statistics in flow cytometry.
Gives the **average intensity** of a parameter in a population.

Two types :

➔ the **arithmetic** mean

➔ the **geometric** mean.

Choosing the wrong one can impact the results.

Cell-Bio

Some definitions

- Arithmetic Mean ("average")
  - Sum of the "n" individual values of a group divided by n

    Arithmetic mean $=(V_1 + V_2 + V_3 \ldots +V_n)/n$

Geometric Mean

  - Multiply the "n" individual values of a cluster together and get the n^{th} root of this product.

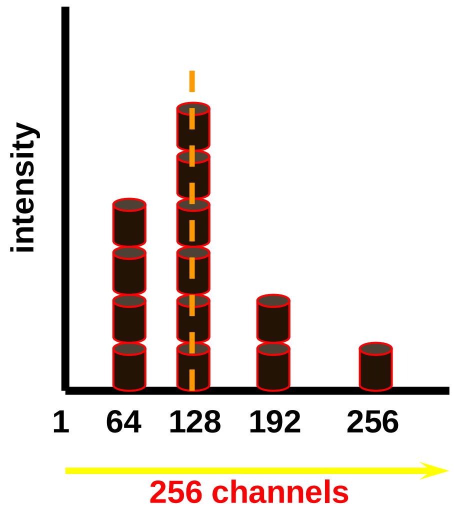
    Geometric mean $= \sqrt[n]{(V_1 \times V_2 \times V_3 \ldots \times V_n)}$

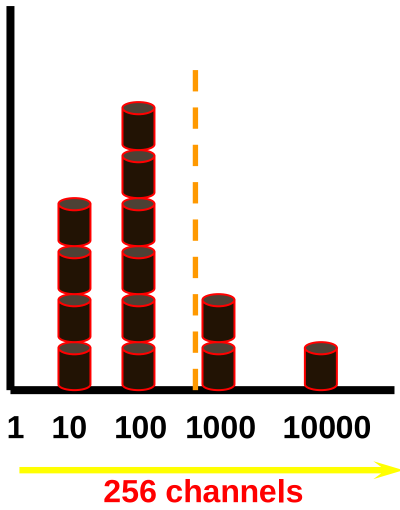
Cell-Bio

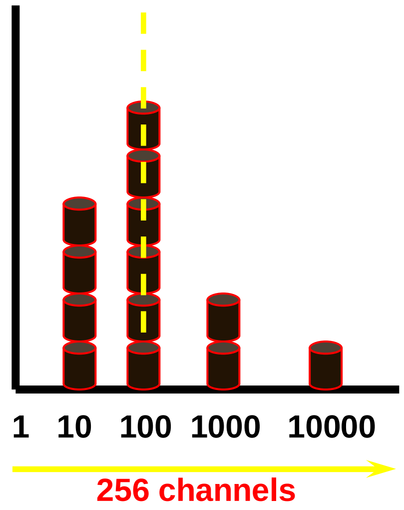
What does it mean?

Linear scale

Logarithmic scale







Arithmetic mean:

$$\frac{4\text{x}64 + 6\text{x}128 + 2\text{x}192 + 256\text{x}1}{13} = \mathbf{128}$$

Arithmetic mean:

$$\frac{4\text{x}10 + 6\text{x}100 + 2\text{x}1000 + 10000\text{x}1}{13} = \mathbf{972.30}$$

→ **NOT display resolution dependent**

Sensitive to small numbers of events in the higher decades

Geometric mean:

$$\sqrt[13]{10^4 \text{ x}100^6 \text{ x } 1000^2 \text{ x } 10000^1} = \mathbf{100}$$

→ **Display resolution dependent**

Cell-Bio

The median

- **Frequently used to describe flow cytometry data.**
- **Refers to the point at which 50% of the events are on either side of a particular channel. *Example : the 2501st cell in a population of 5001.***
- **If population normally distributed : Median = Mean = Mode**
- **Median shifted to a higher intensity value than the mode if the population distribution is skewed to the right and shifted to a lower intensity if skewed to the left.**

If data pile up in the last channel, how far off scale are they ?

→ Impossible to get a true mean value

→Median gives a better information about the central tendency of the population

→ If more than half the population is off-scale, then median and mean cannot give the central tendency of the population.

Other Statistics

Standard Deviation (Sd)

Measures the spread of a distribution
= the dispersion of the values from each event around the mean of a population.

Coefficient of Variation

Defined as the (Standard Deviation /mean) X100.
➔ CVs are always a percentage
➔ Measure of the peak width.

Mode

The mode is the most frequently occurring value in a data range.
If symmetrical distribution, then mode = mean = median
If the distribution is skewed, then these three values are different.

Skewness

Characterizes the asymmetry of a distribution ➔ So it is related to the mean value of the population.
If Value < 0 ➔ asymmetrical distribution ➔ tail towards the left ➔ lower values with respect to the mean.
If Value > 0 ➔ tail towards the right ➔ higher values with respect to the mean.

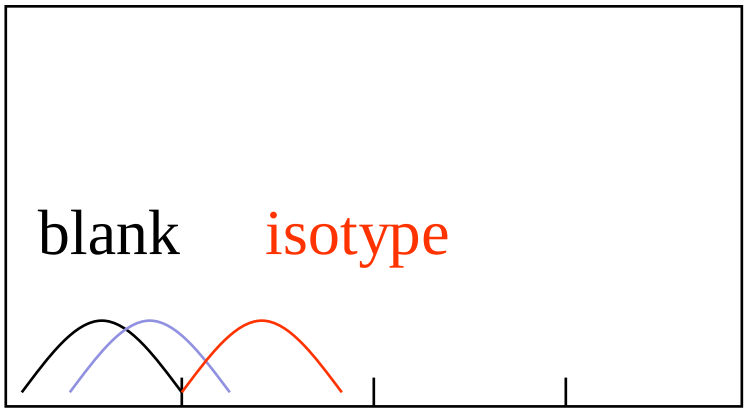
Kurtosis

Kurtosis refers to the relative "flatness" of a distribution and is also related to the mean of the distribution.
A Value<0 ➔ relatively flat distribution,
A Value>0 ➔ a relatively peaked distribution
} compared to the normal distribution

Cell-Bio

Control is Important!

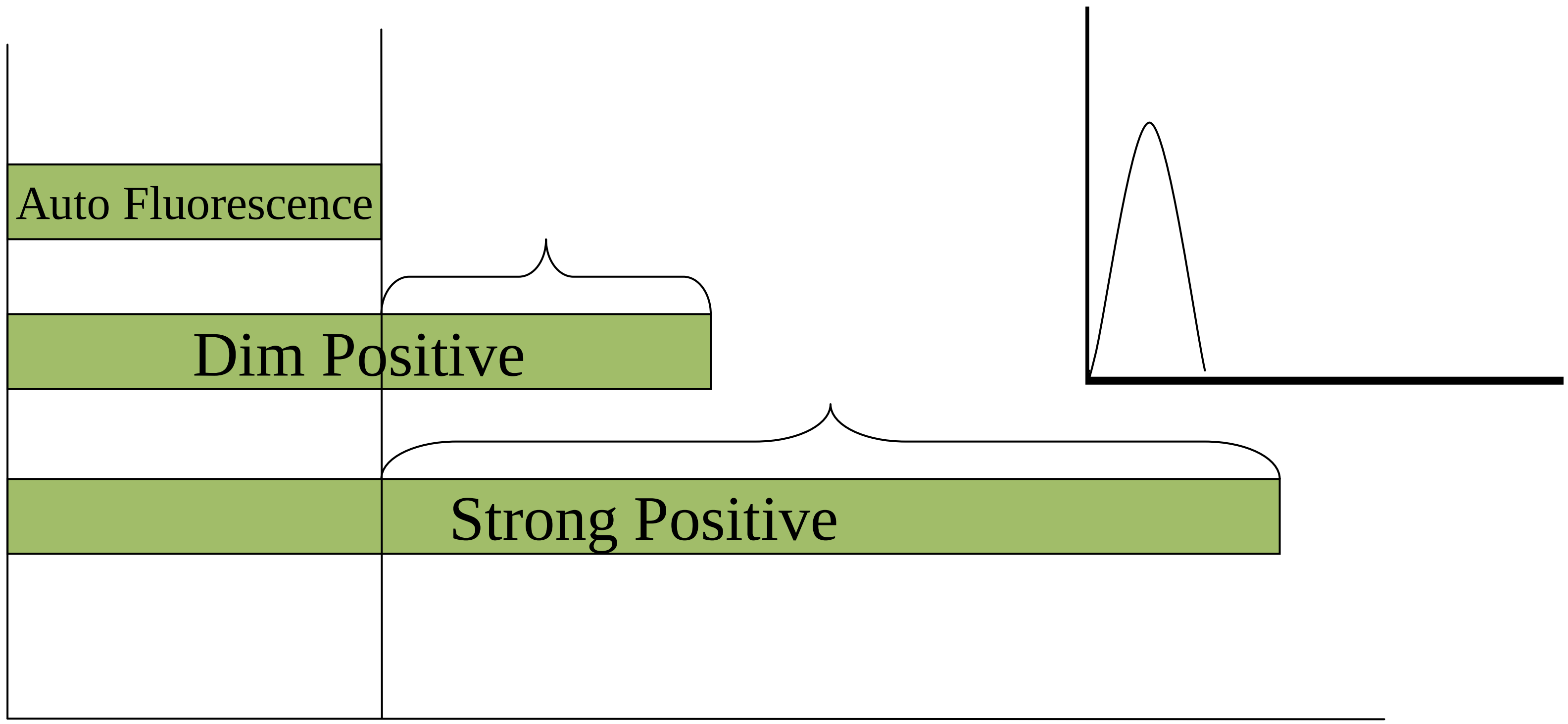
- Blank Control: autofluorescence, instr setup.
- Negative Control: the extent of non-specific staining
- Isotype Control: for indirect staining
- Positive Control: antibody functionality



Control is Important!

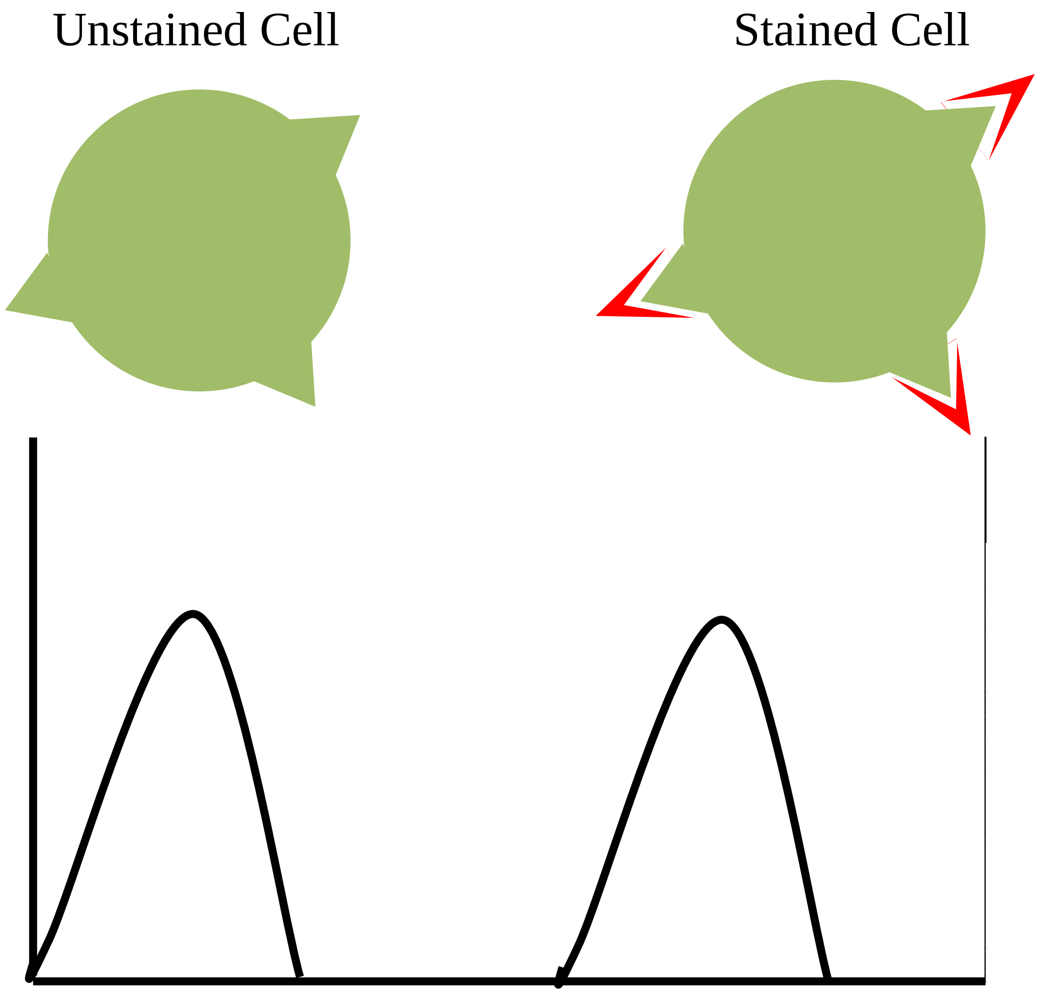
- Blank Control: use Untreated or Unstained Cell to distinguish auto fluorescence.

Why Blank Cell??



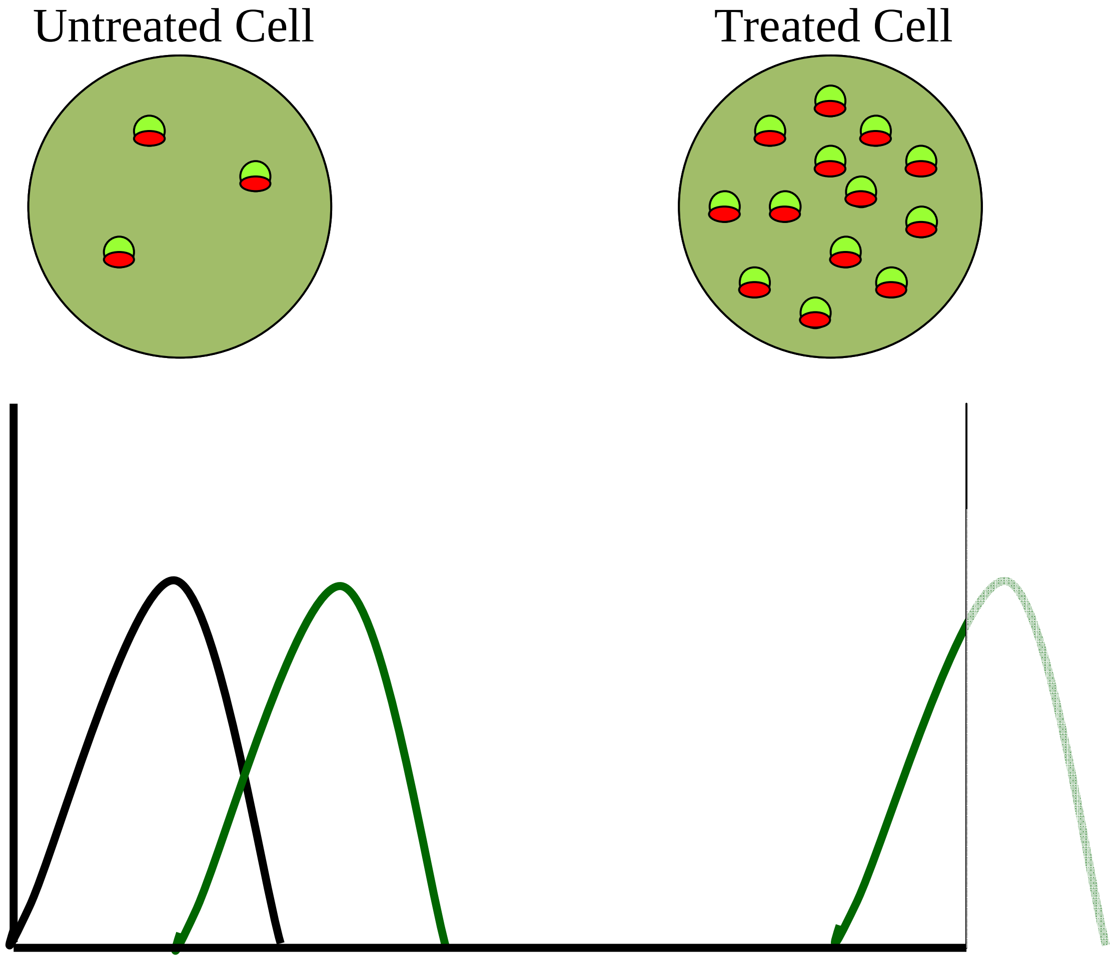
Control is Important!

- Blank Control: use Untreated or Unstained Cell to distinguish auto fluorescence.



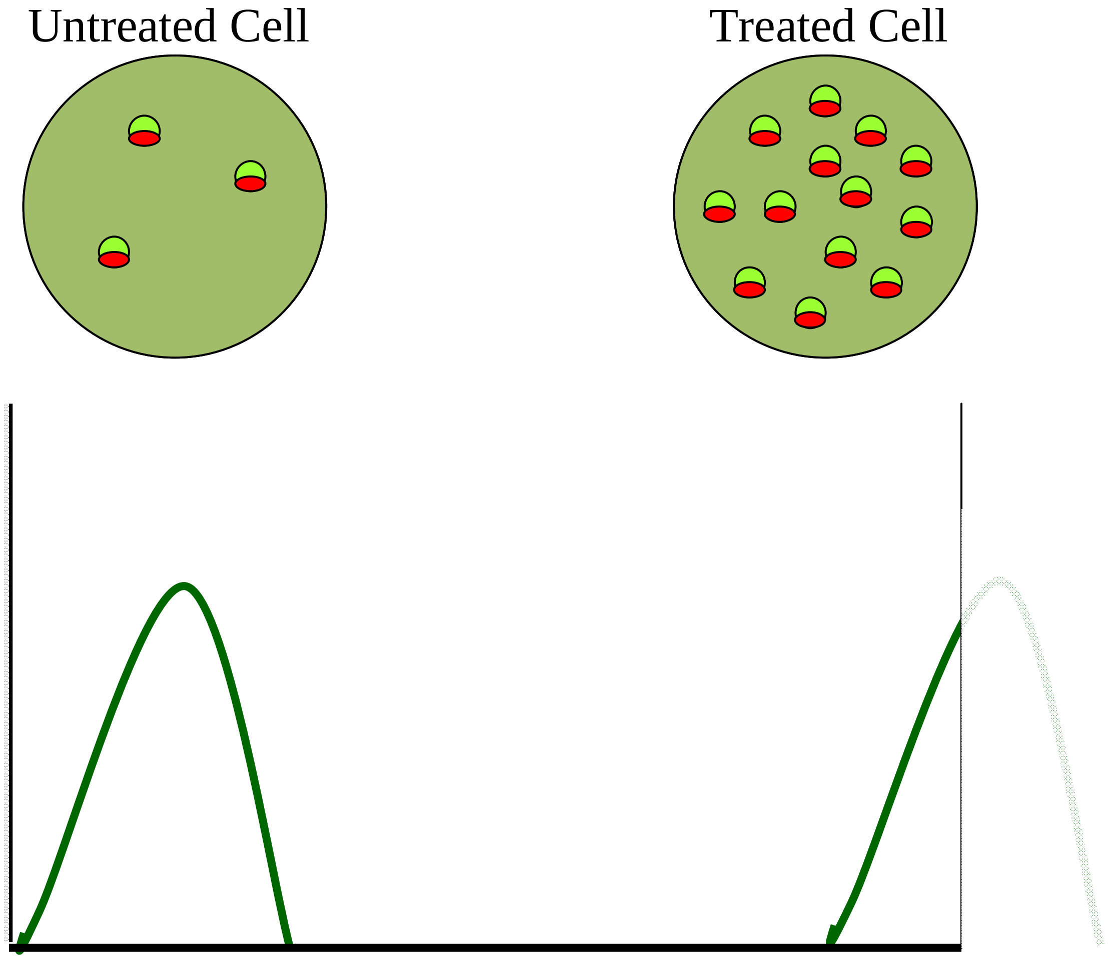
Control is Important!

- Blank Control: use Untreated or Unstained Cell to distinguish auto fluorescence.



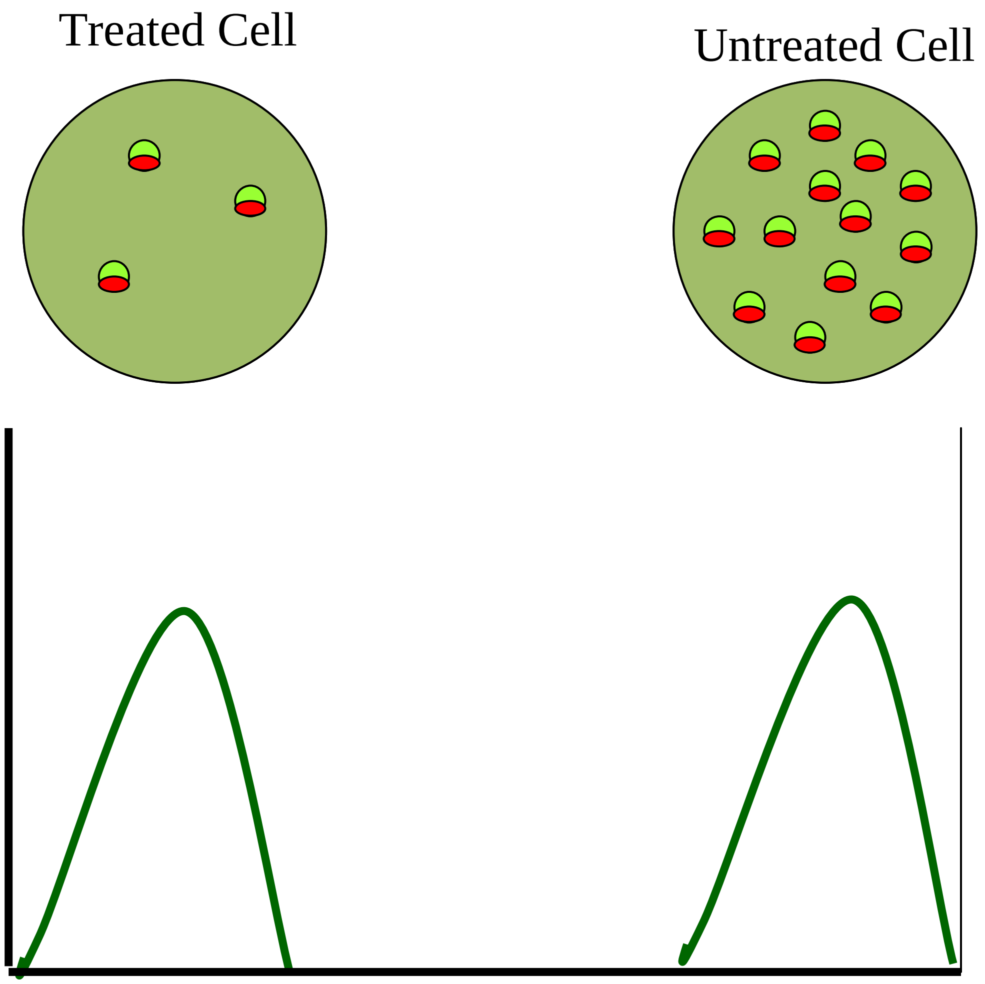
Control is Important!

- Blank Control: use Untreated or Unstained Cell to distinguish auto fluorescence.



Control is Important!

- Inverted Control.



陽性界限設定原則

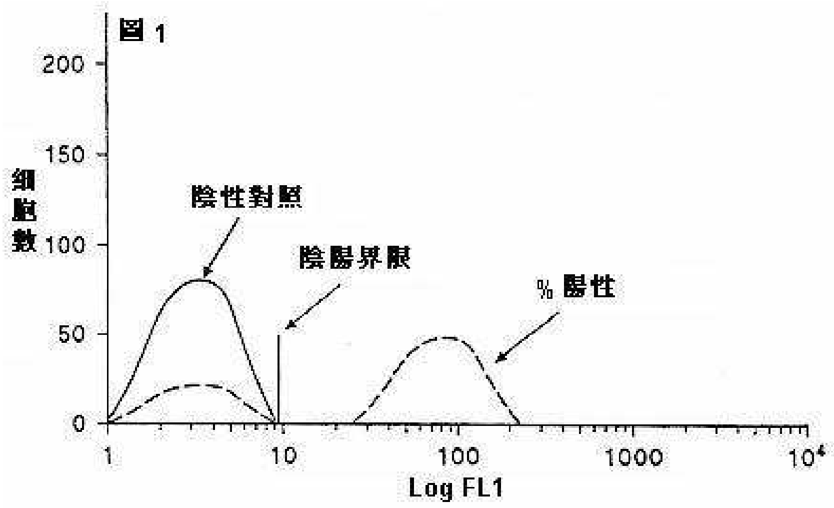


圖 1
200
150
100
50
0
細胞數
陰性對照
陰陽界限
% 陽性
1
10
100
1000
10^4
Log FL1

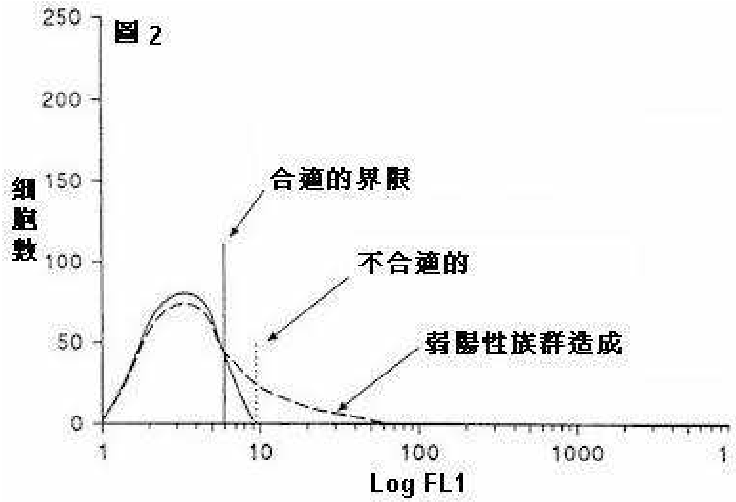


圖 2
250
200
150
100
50
0
細胞數
合適的界限
不合適的
弱陽性族群造成
1
10
100
1000
1
Log FL1

陽性界限設定原則

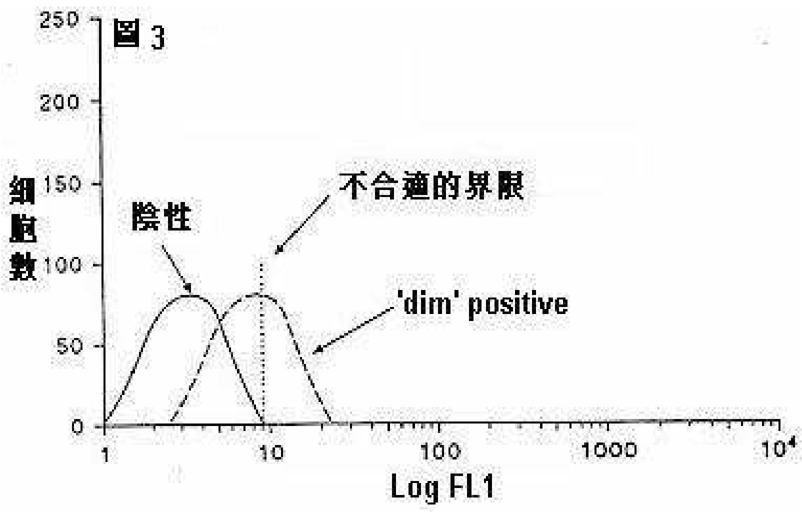


圖 3
細胞數
陰性
不合適的界限
'dim' positive
Log FL1

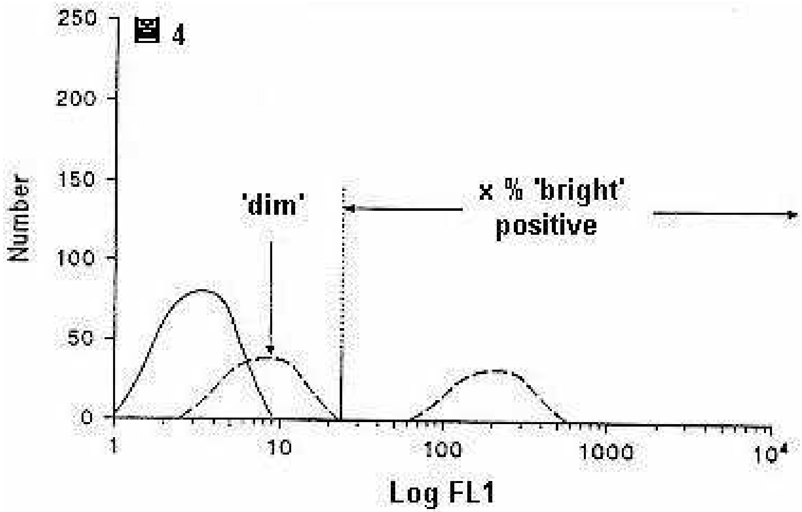


圖 4
Number
'dim'
x % 'bright' positive
Log FL1