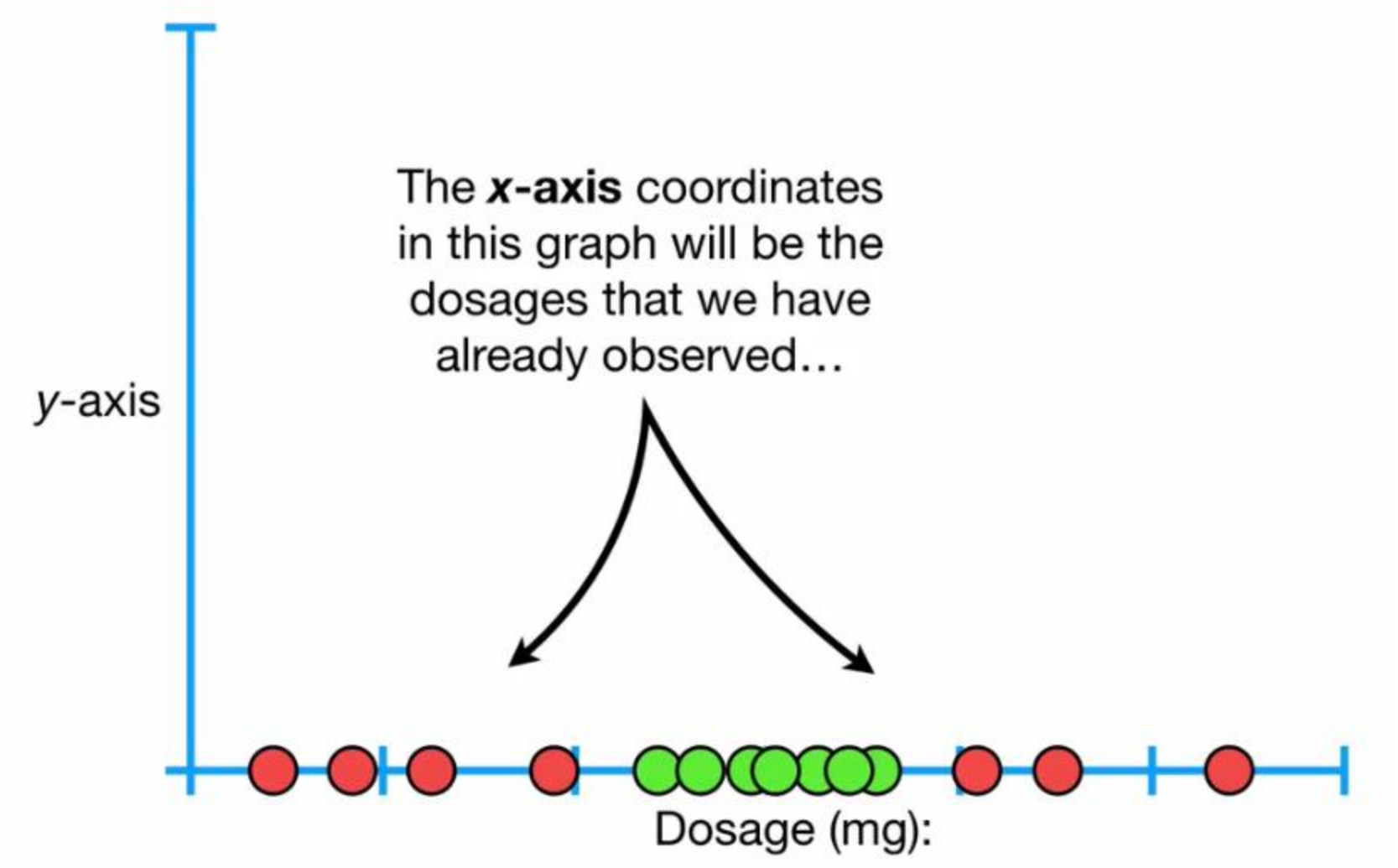
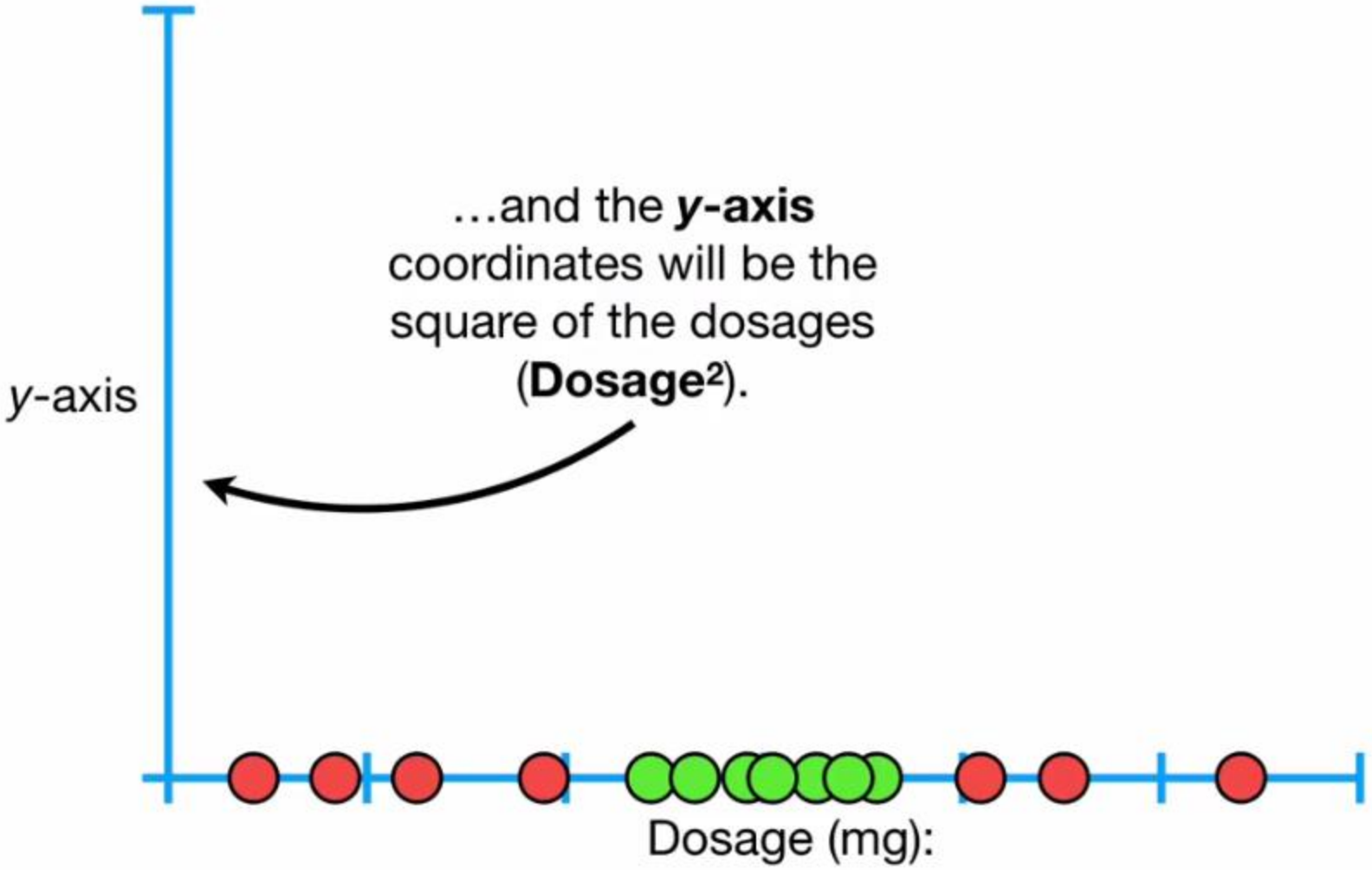
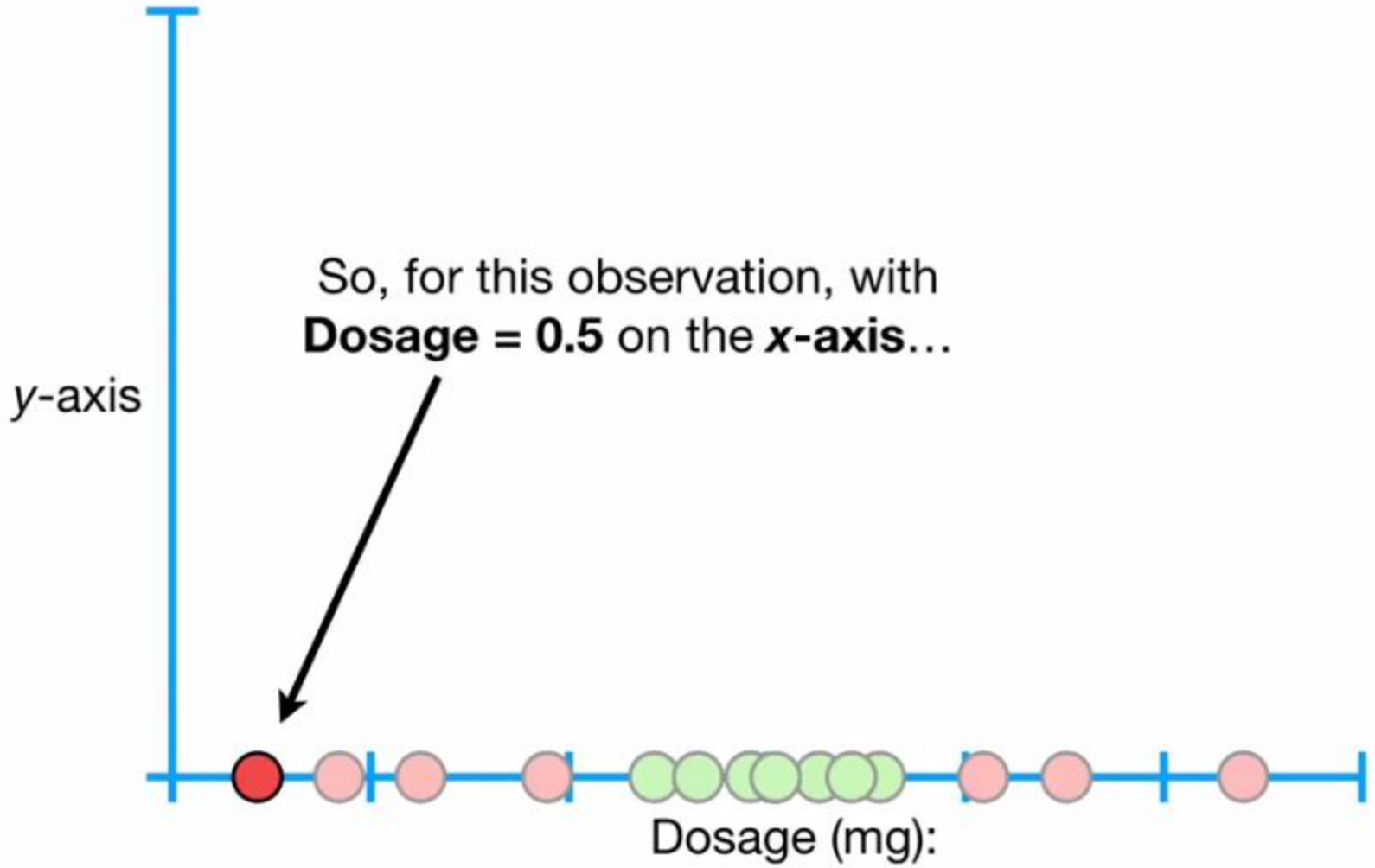
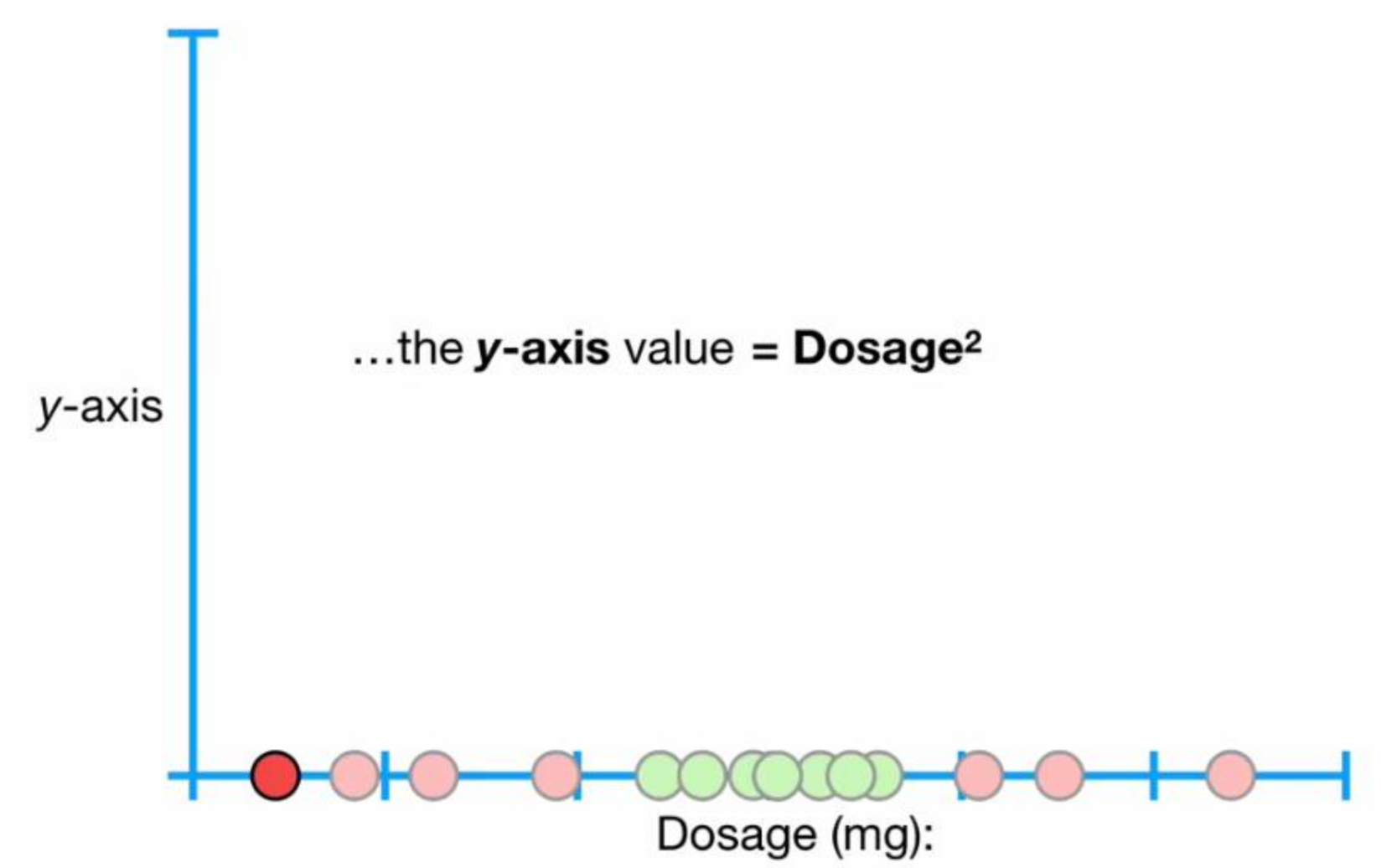
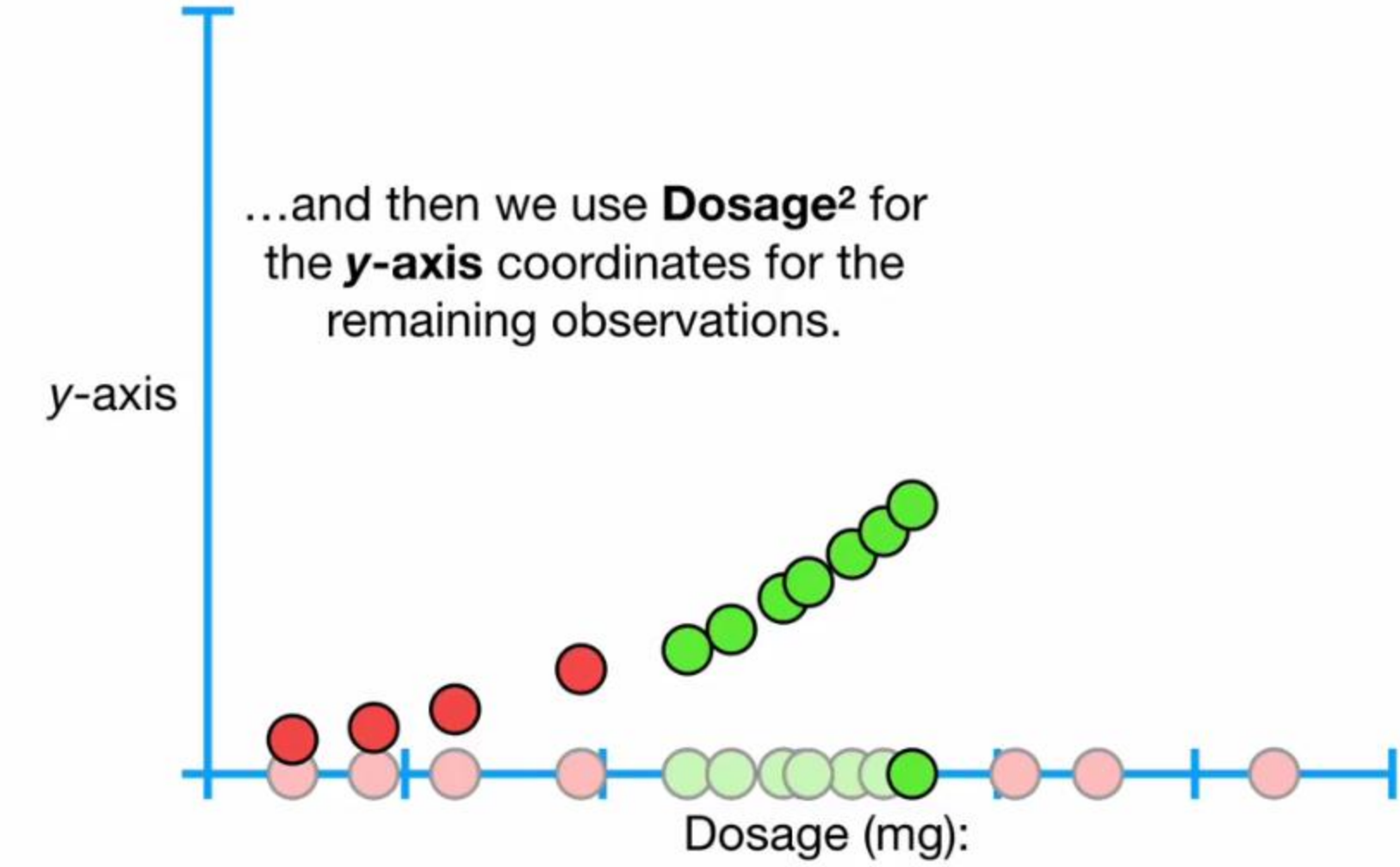
Support Vector Machine

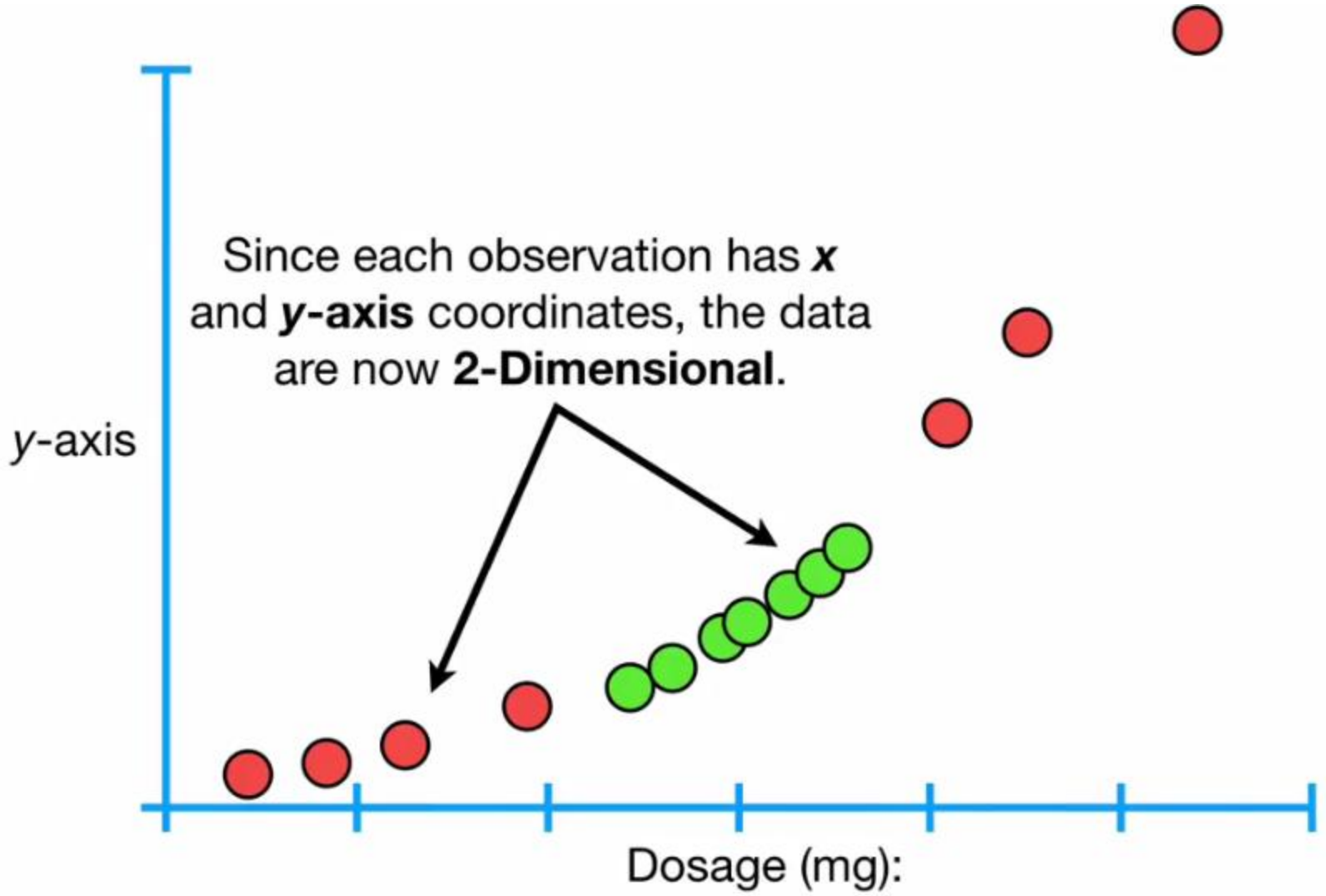

The x-axis coordinates in this graph will be the dosages that we have already observed…
y-axis
Dosage (mg):

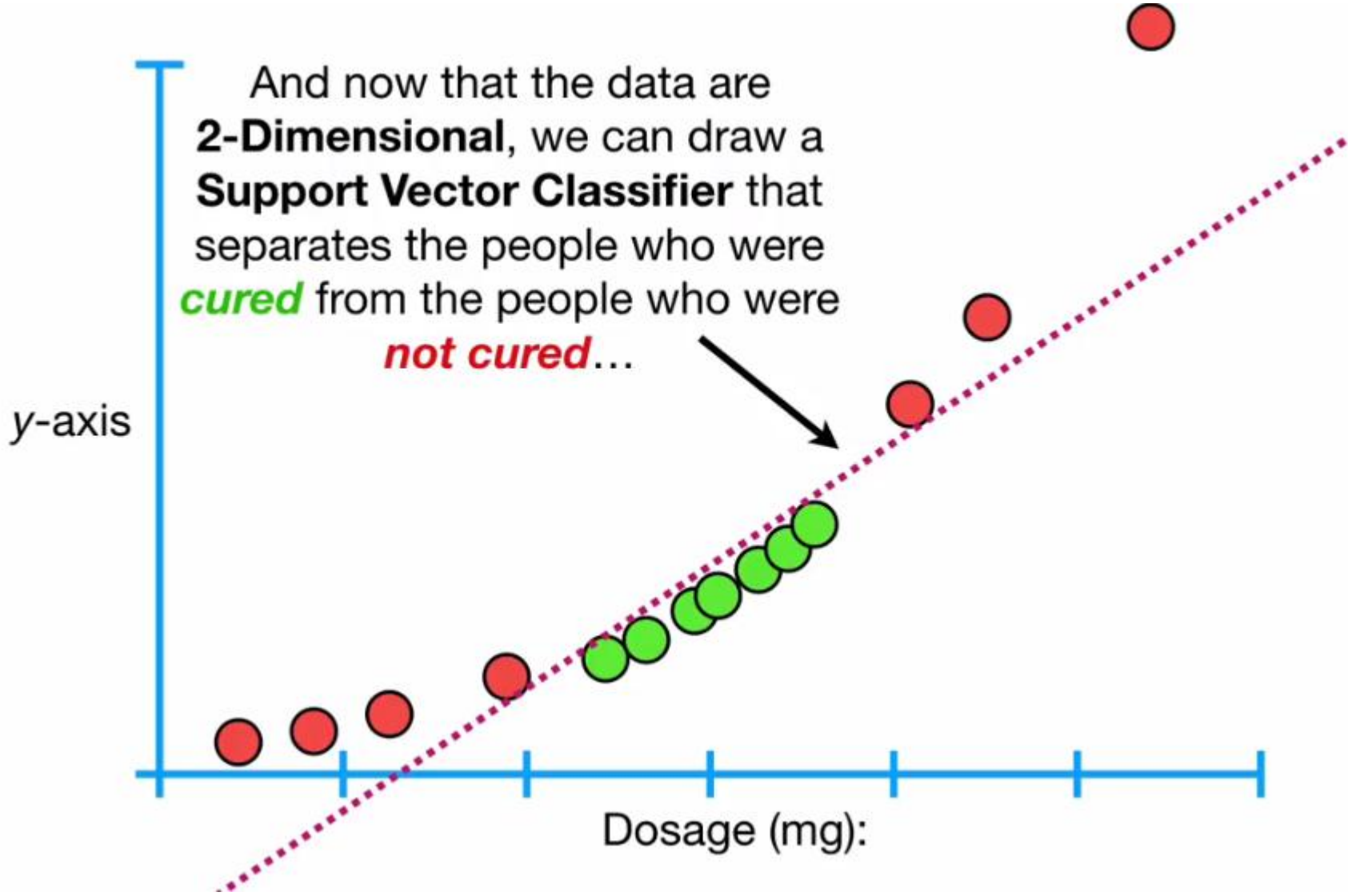

…and the **y-axis** coordinates will be the square of the dosages (**Dosage²**).
y-axis
Dosage (mg):

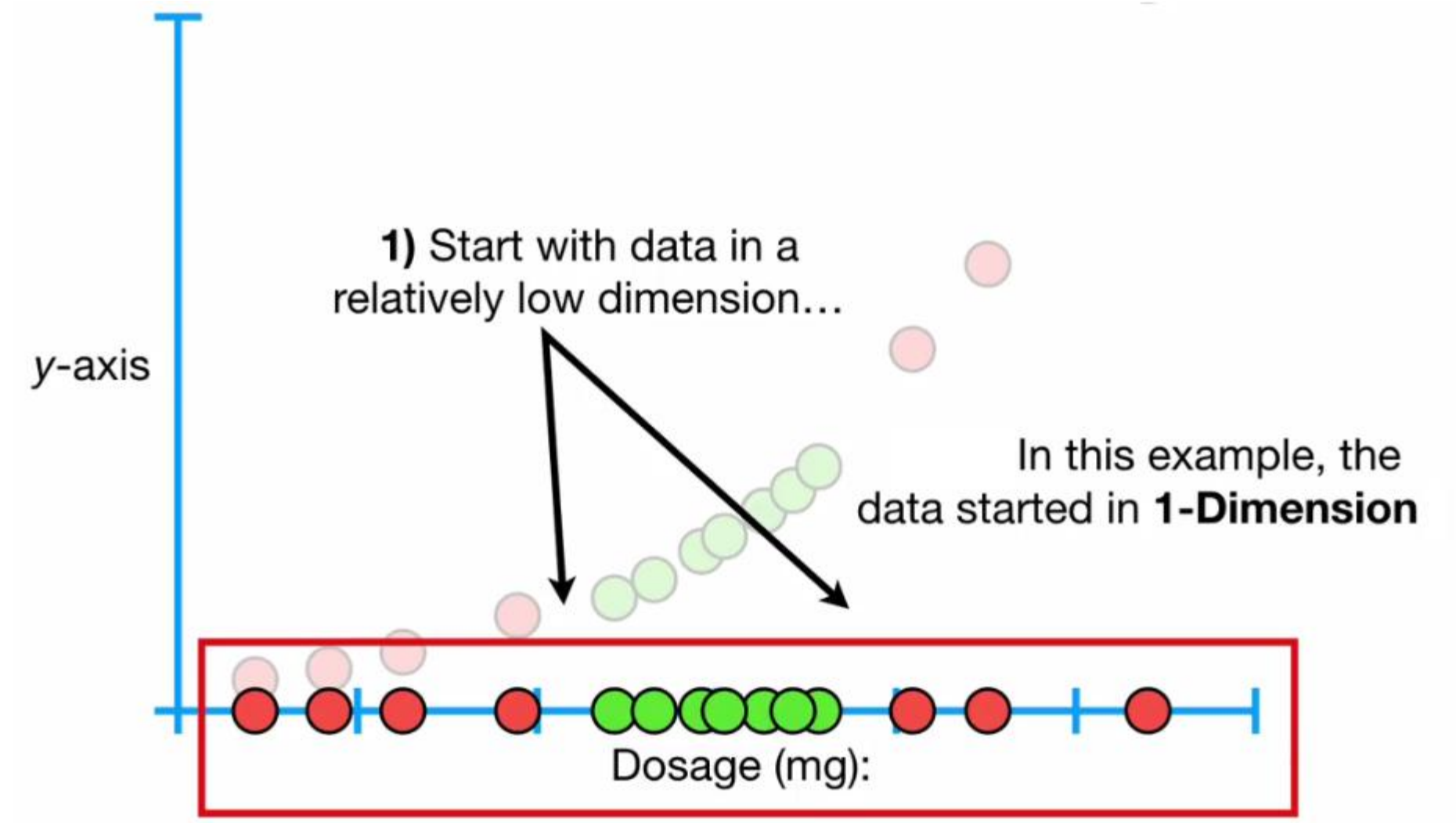

So, for this observation, with
Dosage = 0.5 on the ***x*-axis**…
y-axis
Dosage (mg):

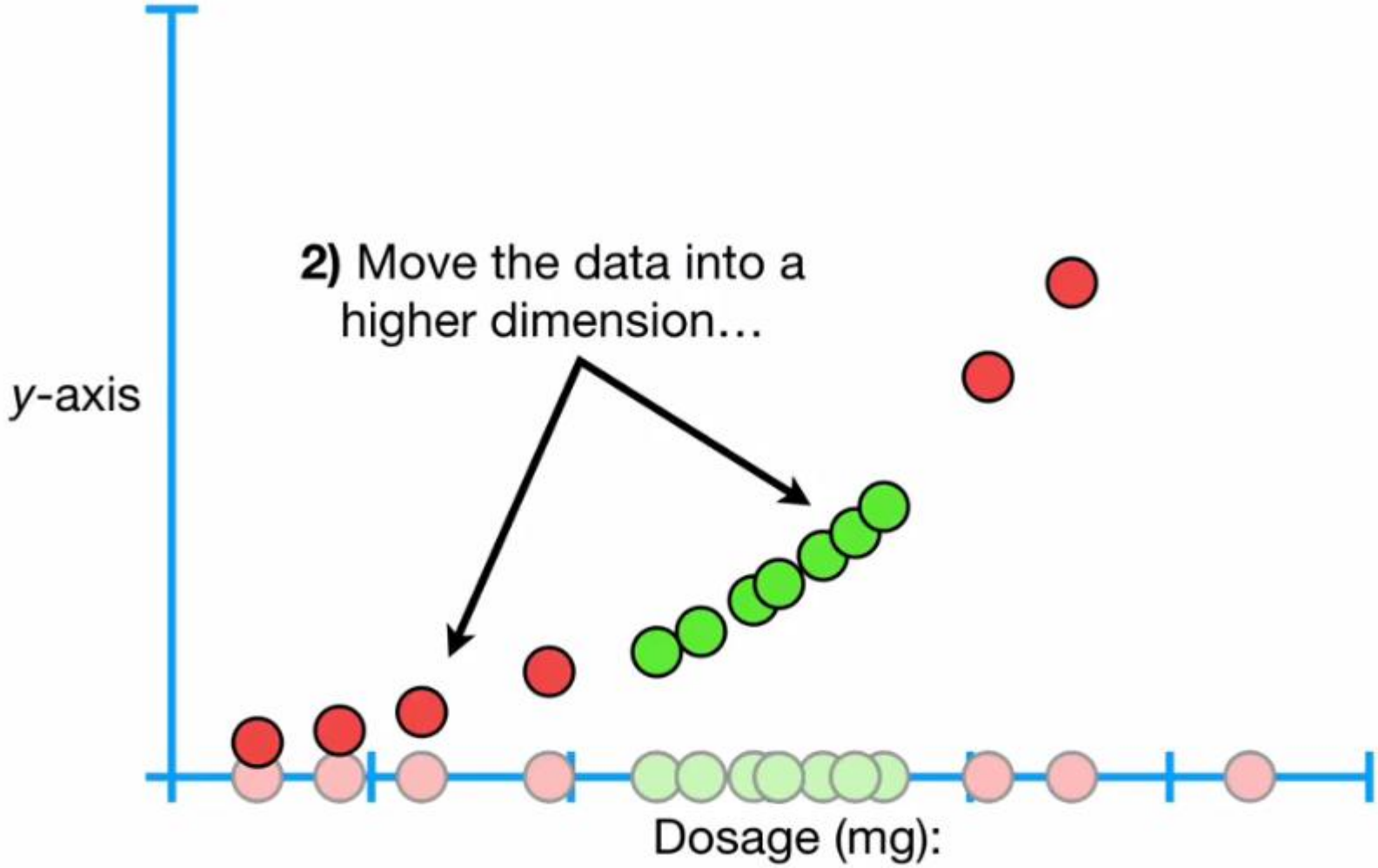

…the **y-axis** value = **Dosage²**
y-axis
Dosage (mg):

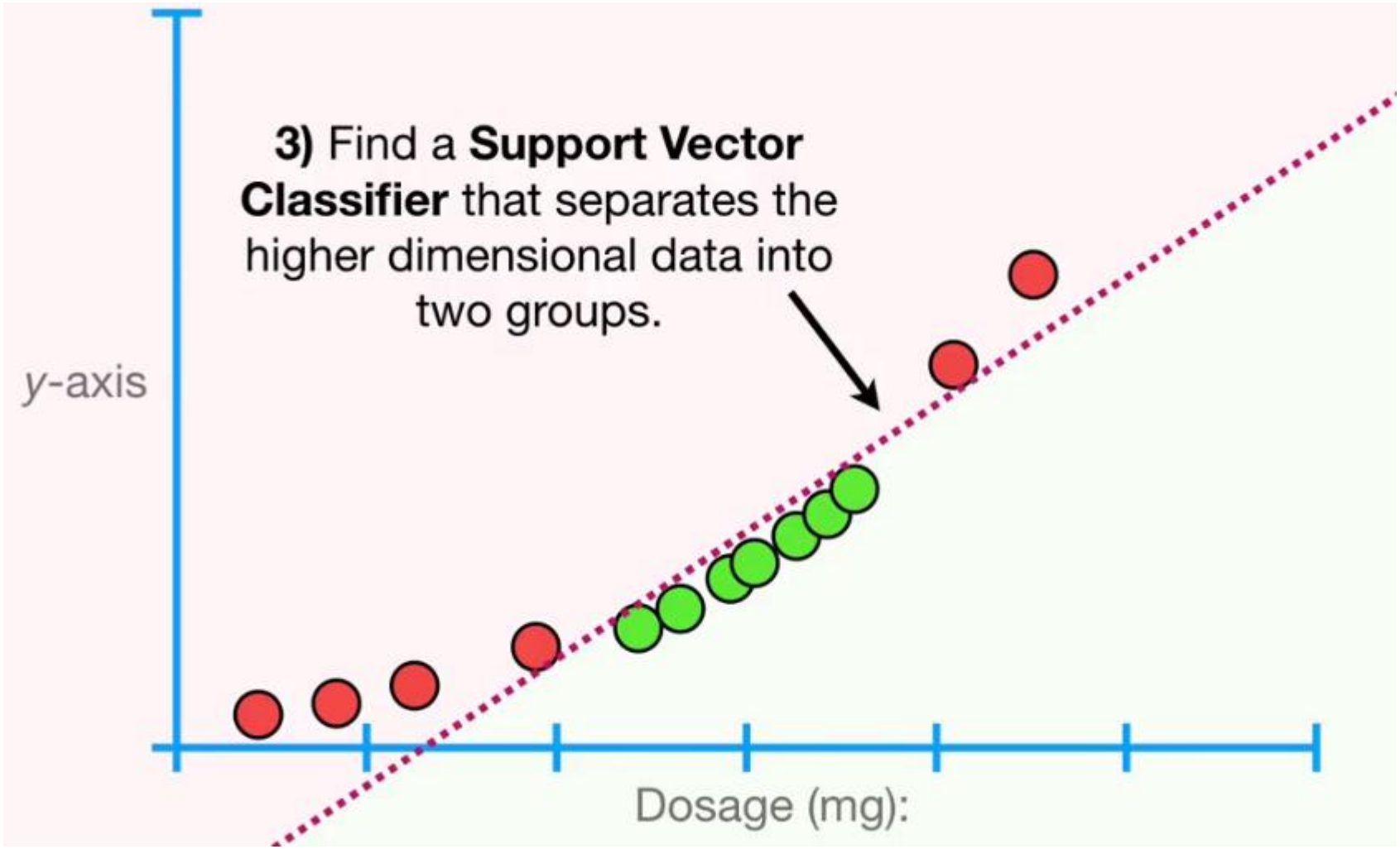

…and then we use **Dosage²** for the ***y*-axis** coordinates for the remaining observations.
y-axis
Dosage (mg):

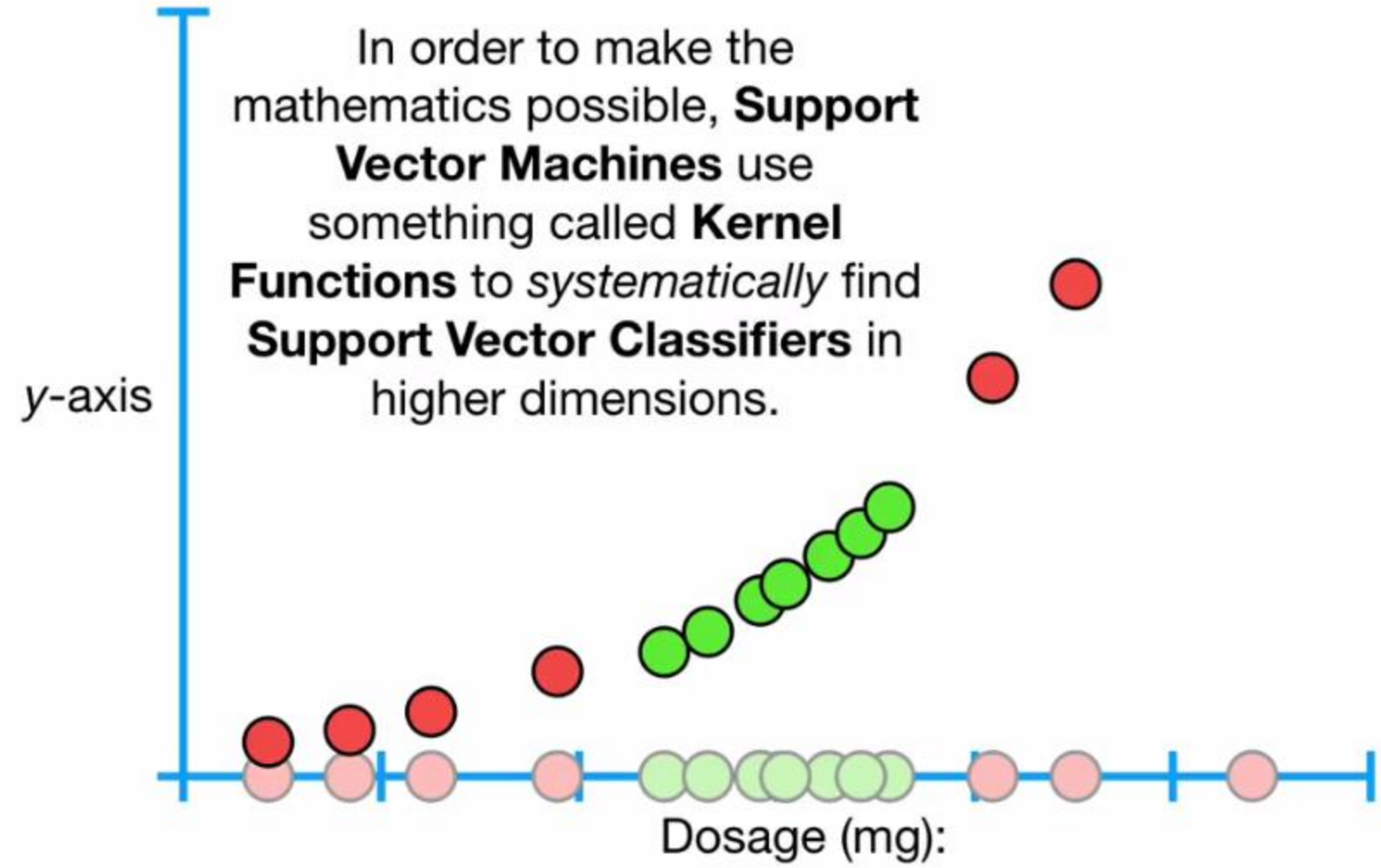

Since each observation has *x* and ***y*-axis** coordinates, the data are now **2-Dimensional**.
y-axis
Dosage (mg):

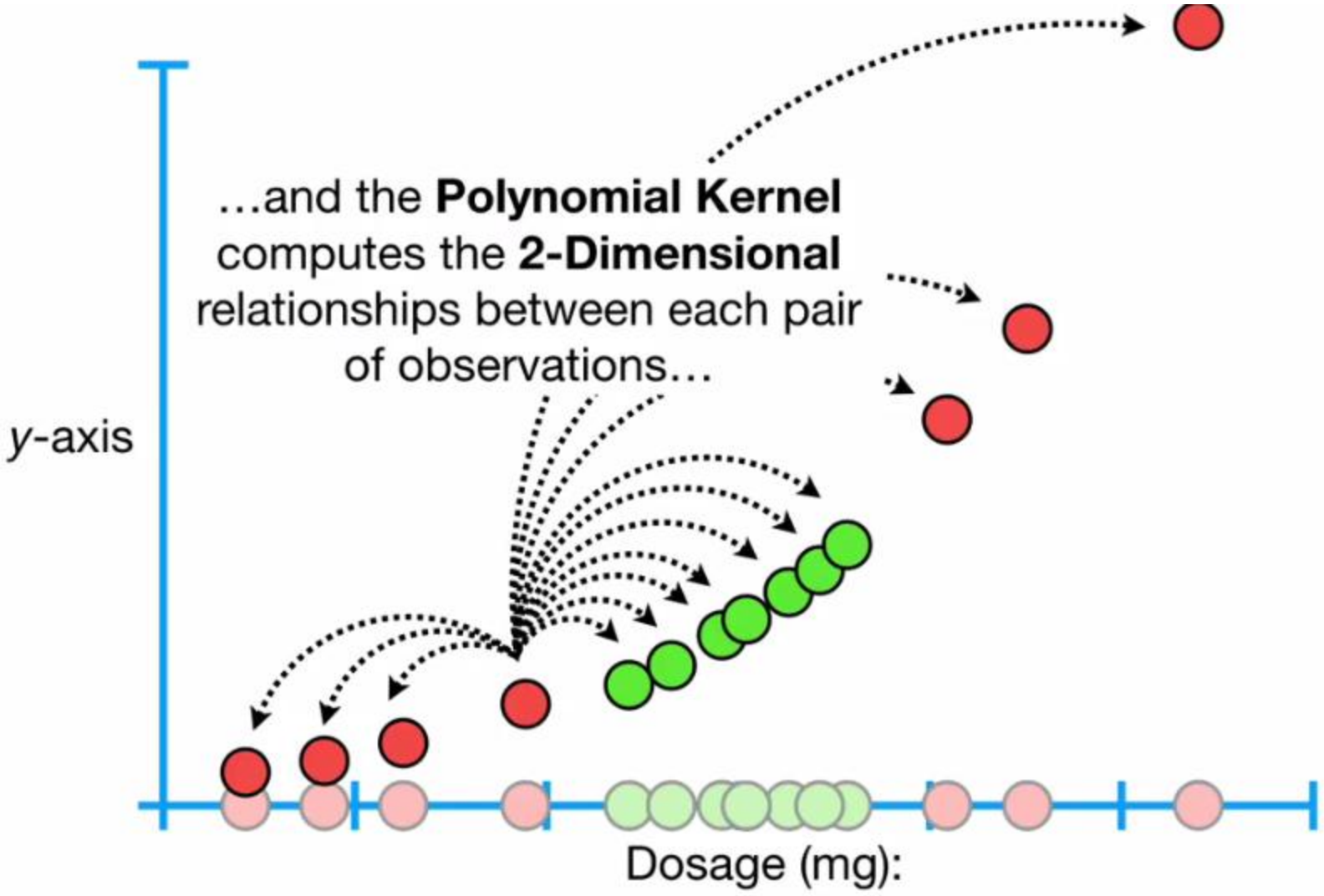

And now that the data are
2-Dimensional, we can draw a
Support Vector Classifier that
separates the people who were
cured from the people who were
not cured…
y-axis
Dosage (mg):


y-axis
1) Start with data in a relatively low dimension…
In this example, the data started in 1-Dimension
Dosage (mg):


2) Move the data into a higher dimension…
y-axis
Dosage (mg):

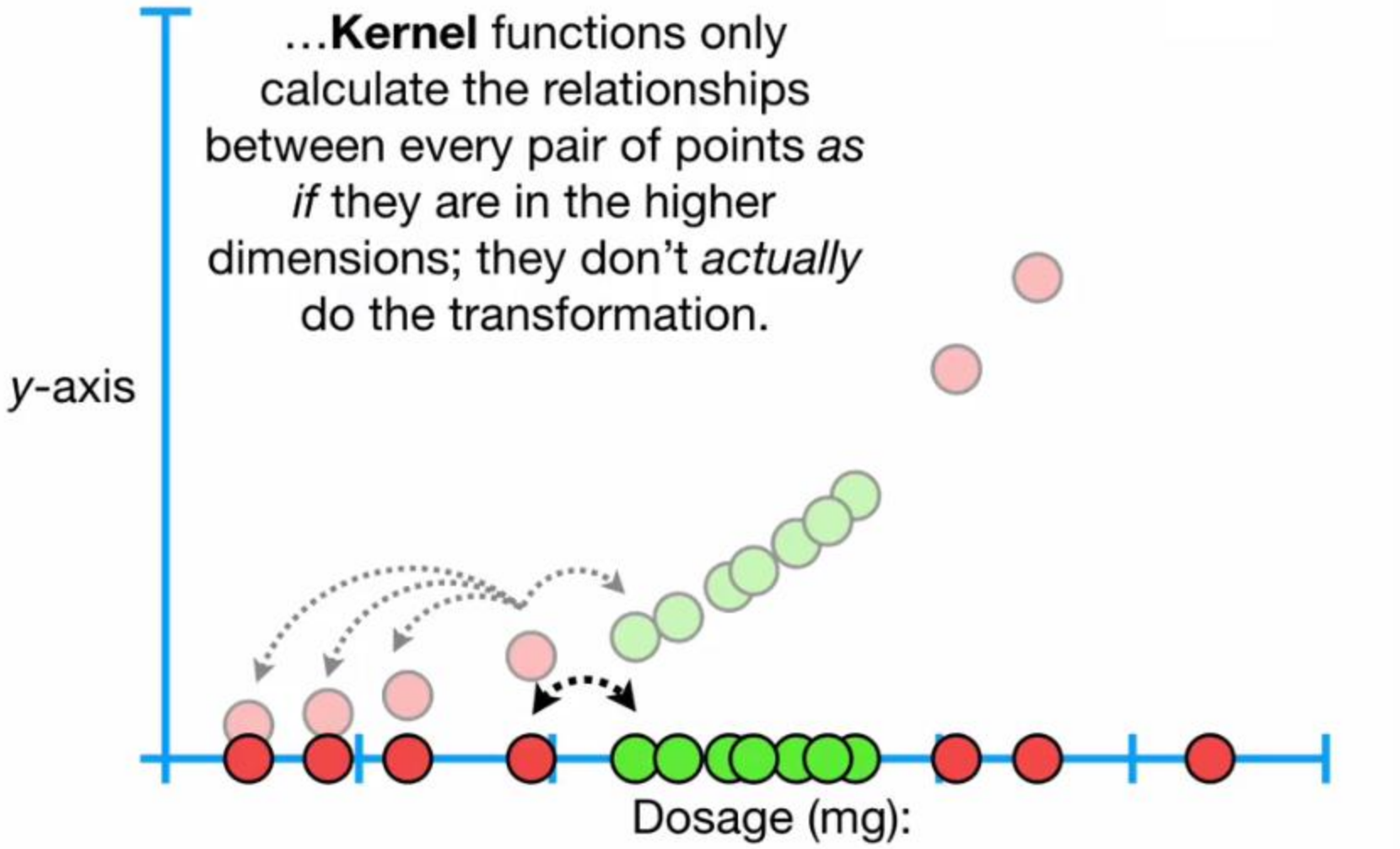

3) Find a **Support Vector Classifier** that separates the higher dimensional data into two groups.
y-axis
Dosage (mg):

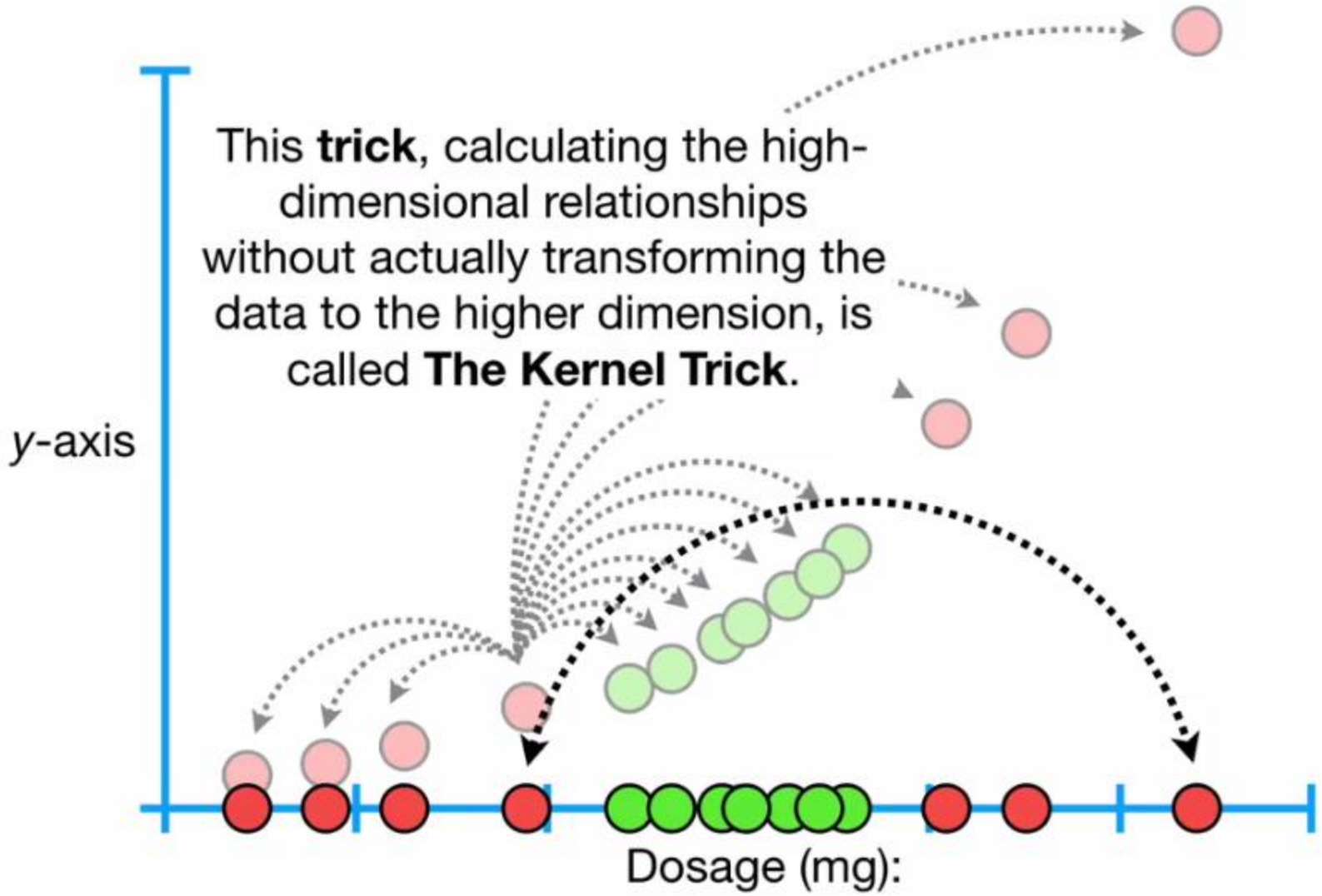

In order to make the mathematics possible, **Support Vector Machines** use something called **Kernel Functions** to *systematically* find **Support Vector Classifiers** in higher dimensions.
y-axis
Dosage (mg):

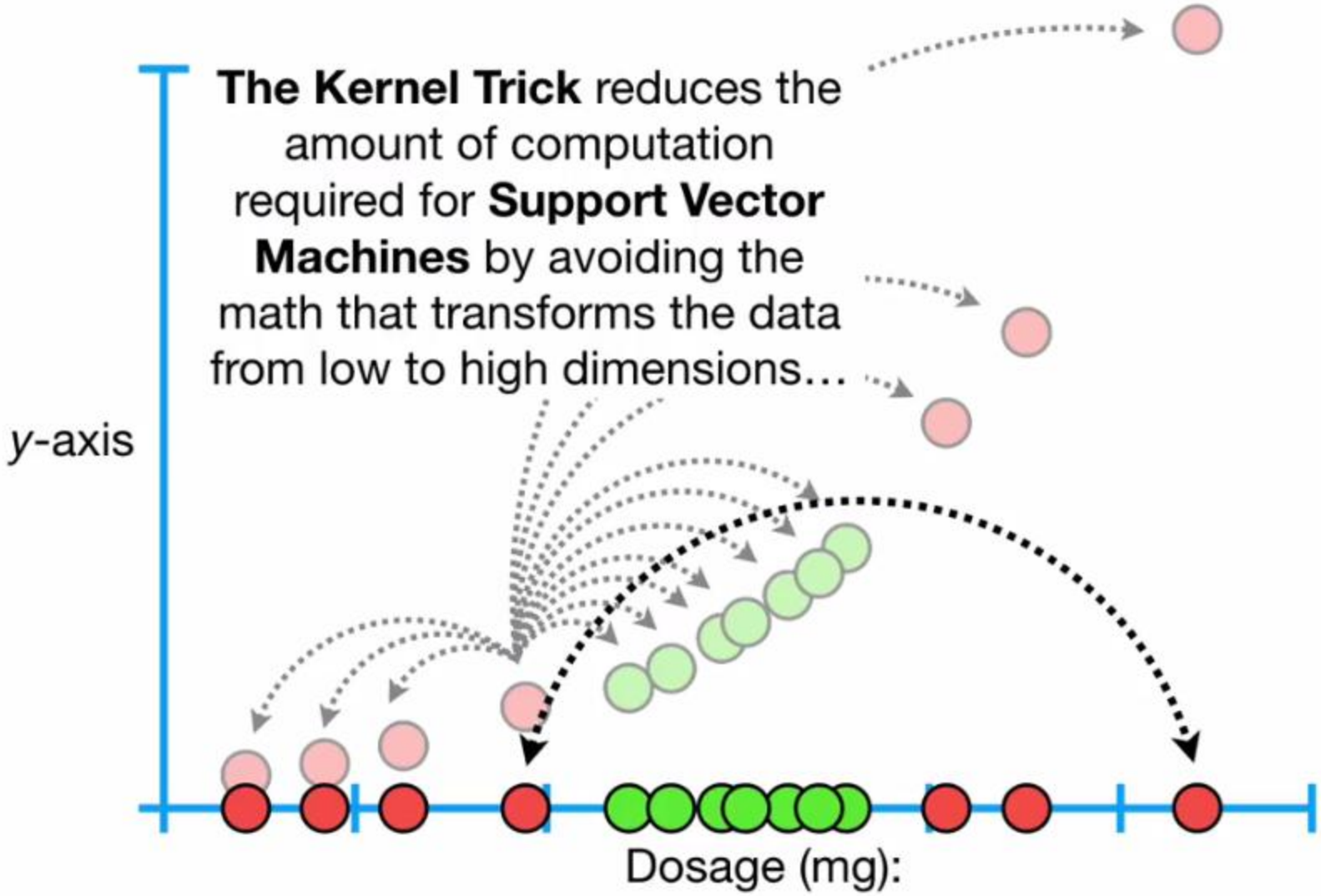

…and the **Polynomial Kernel** computes the **2-Dimensional** relationships between each pair of observations…
y-axis
Dosage (mg):

Another very commonly used **Kernel** is the **Radial Kernel**, also known as the **Radial Basis Function (RBF) Kernel**.

Radial Kernel finds **Support Vector Classifiers** in *infinite dimensions*


…**Kernel** functions only calculate the relationships between every pair of points *as if* they are in the higher dimensions; they don't *actually* do the transformation.
y-axis
Dosage (mg):


This **trick**, calculating the high-dimensional relationships without actually transforming the data to the higher dimension, is called **The Kernel Trick**.
y-axis
Dosage (mg):


The Kernel Trick reduces the amount of computation required for **Support Vector Machines** by avoiding the math that transforms the data from low to high dimensions…
y-axis
Dosage (mg):