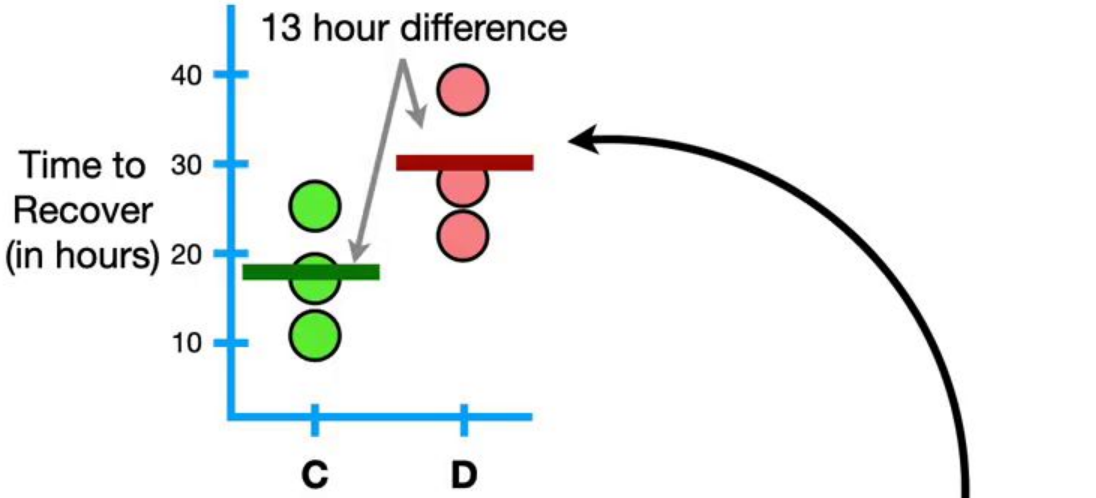
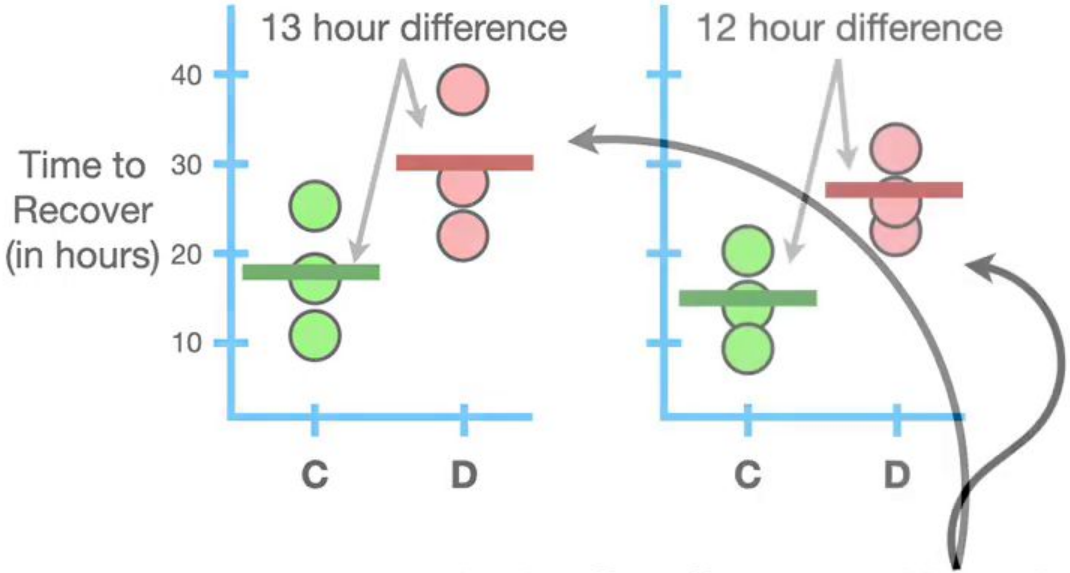
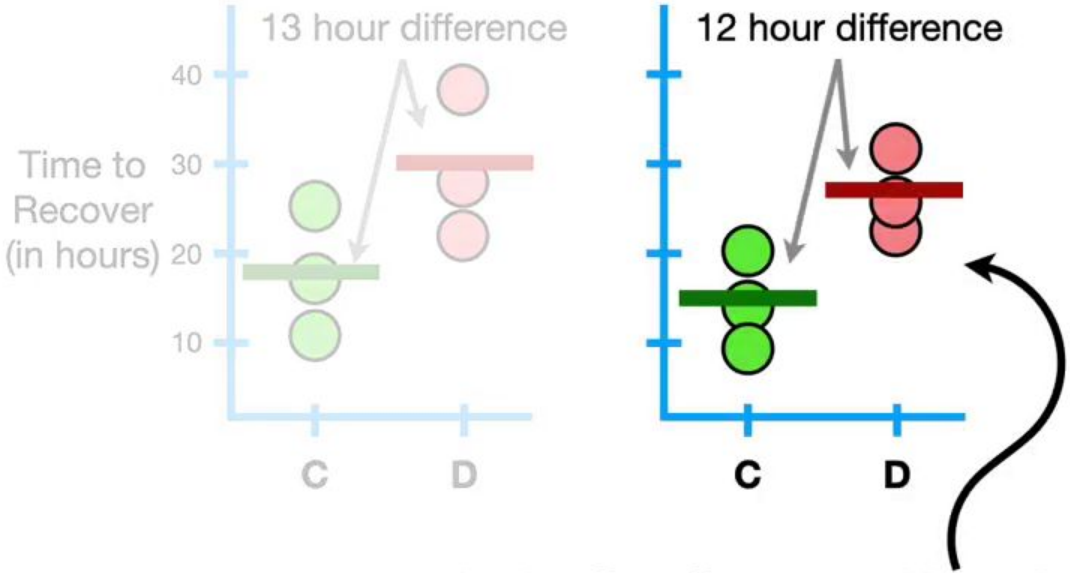
Alternative Hypotheses: Main Ideas!!!



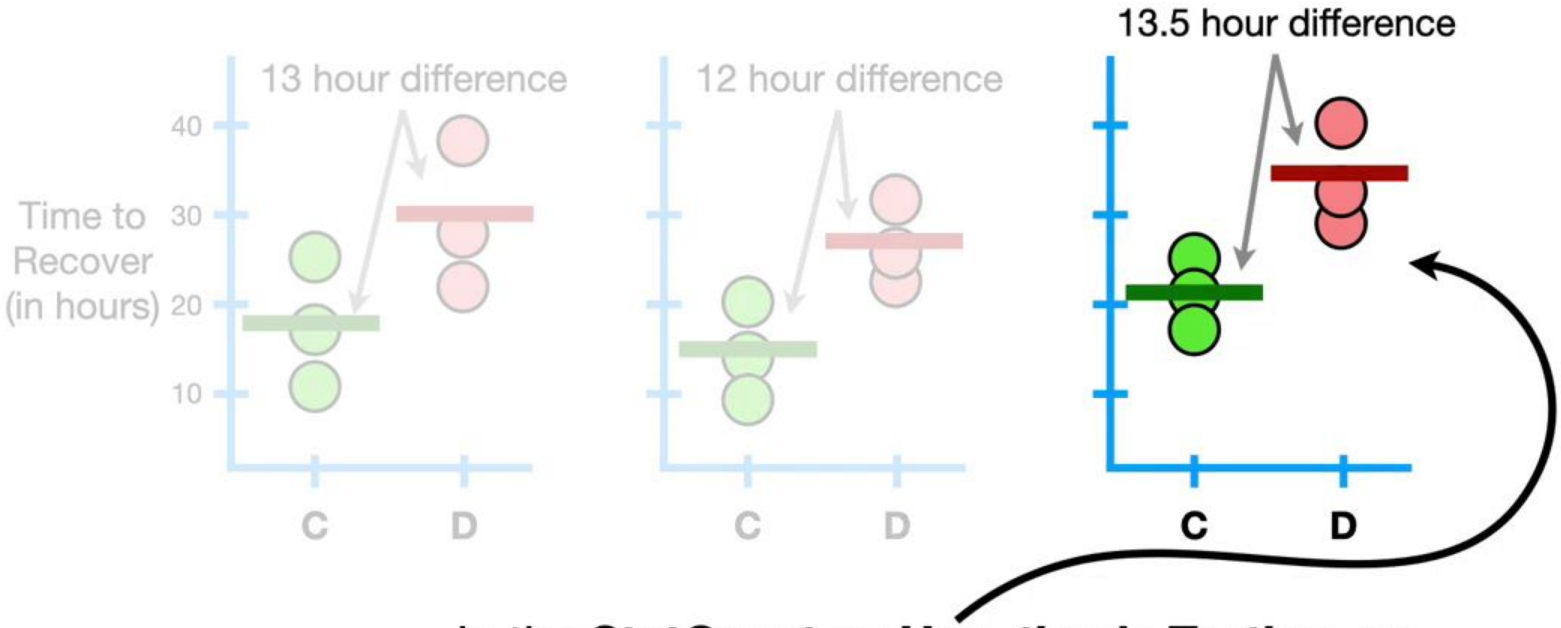
In the **StatQuest** on **Hypothesis Testing**, we learned that rather than get stressed out over a large number of possible hypotheses that we could test to see if two drugs are different…



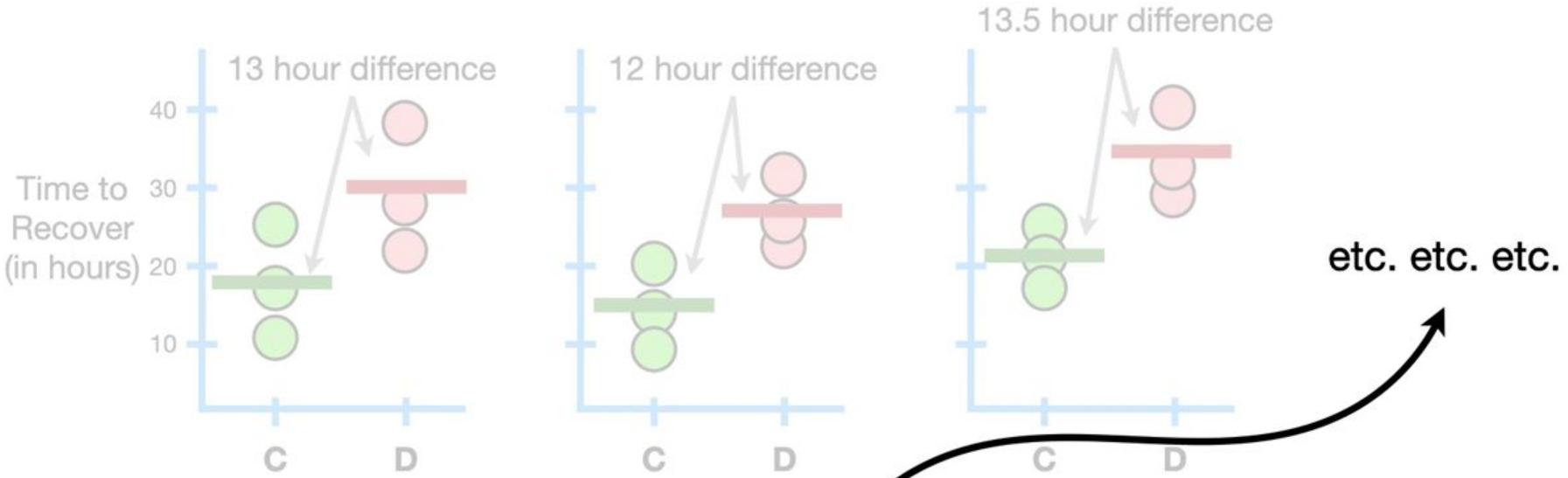
In the **StatQuest** on **Hypothesis Testing**, we learned that rather than get stressed out over a large number of possible hypotheses that we could test to see if two drugs are different…



In the **StatQuest** on **Hypothesis Testing**, we learned that rather than get stressed out over a large number of possible hypotheses that we could test to see if two drugs are different…



In the **StatQuest** on **Hypothesis Testing**, we learned that rather than get stressed out over a large number of possible hypotheses that we could test to see if two drugs are different…

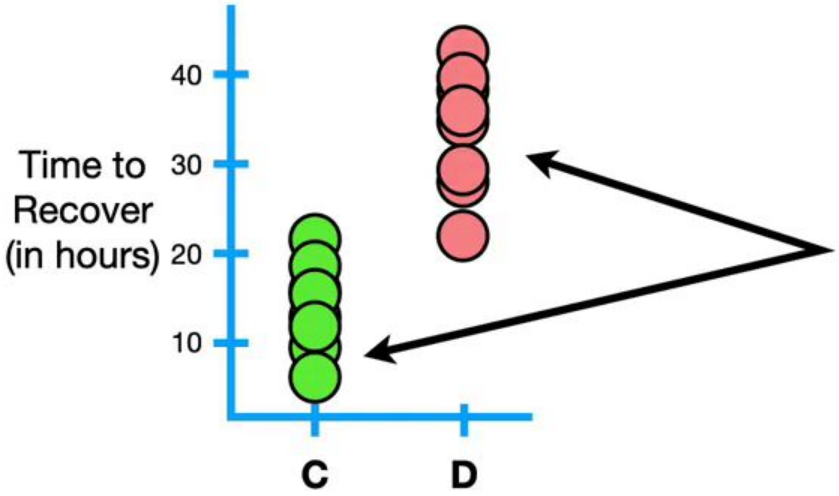


etc. etc. etc.

In the **StatQuest** on **Hypothesis Testing**, we learned that rather than get stressed out over a large number of possible hypotheses that we could test to see if two drugs are different…

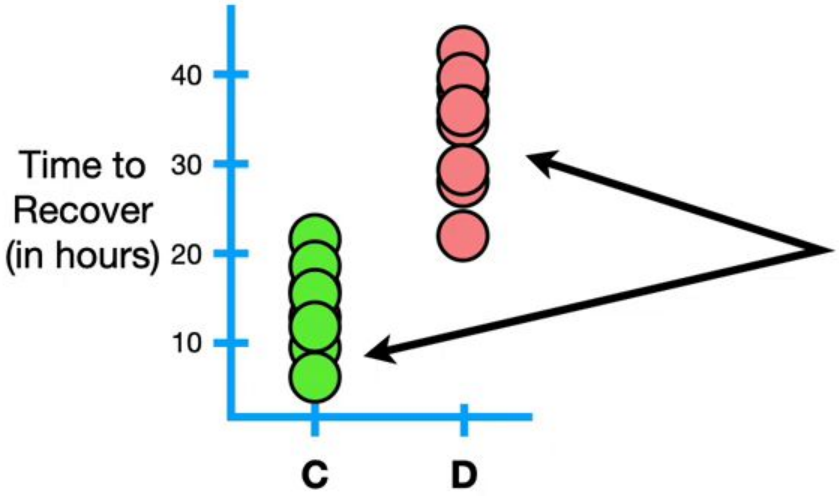
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

…we simply use the **Null Hypothesis** to determine if there is a difference.



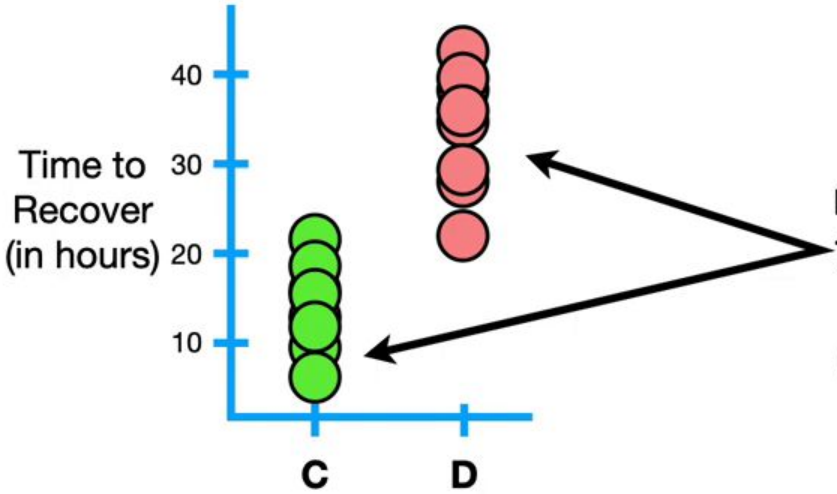
We learned that if we do an experiment with a bunch of people…

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



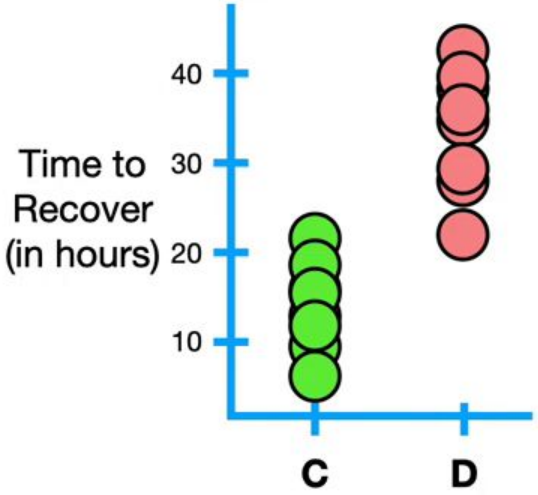
…and a lot more people taking **Drug C** had shorter recovery times than people taking **Drug D**…

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



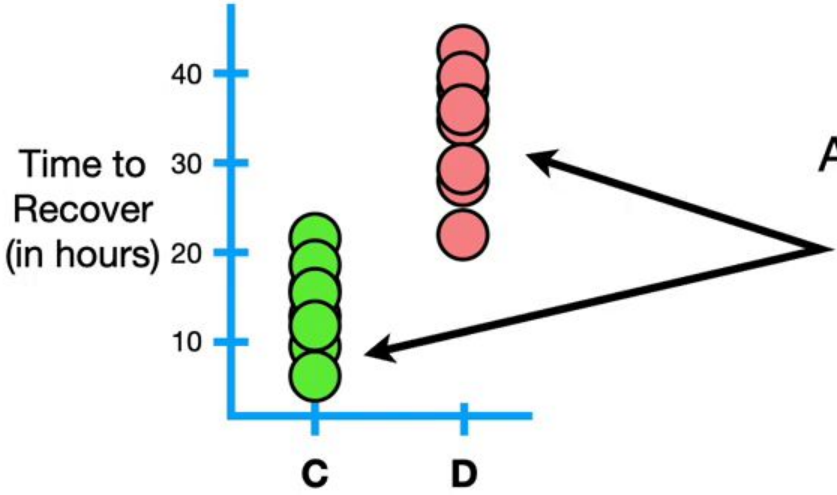
…so many that it would be hard to imagine that the results were due to random things, like everyone taking **Drug C** had better diets or got more exercise than the people taking **Drug D**…

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



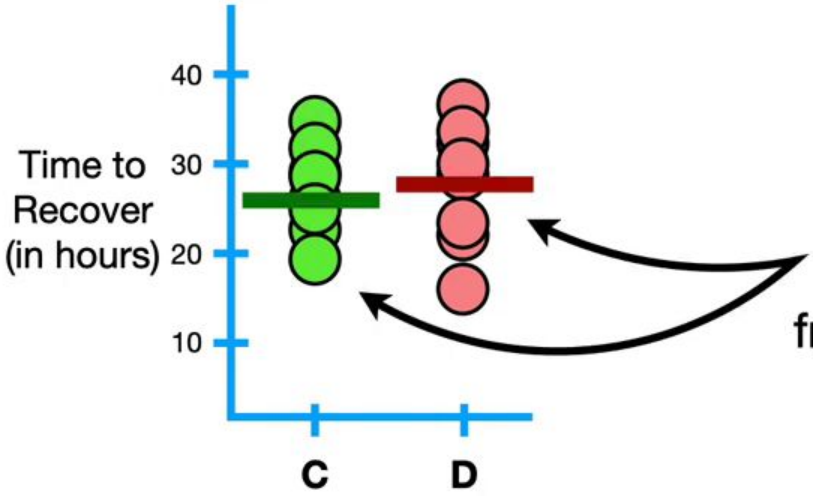
…then we can ***reject the Null Hypothesis***.

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



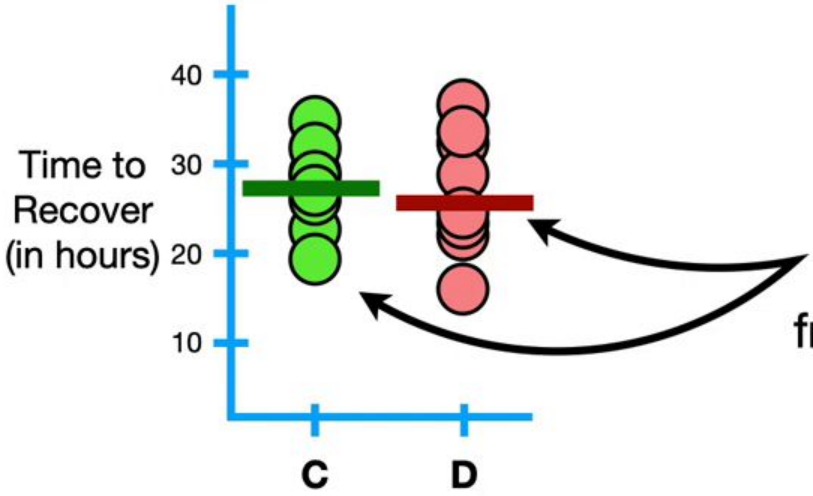
And then we would know that there is difference between **Drug C** and **Drug D**.

~~There is ***no difference*** in recovery times between **Drug C** and **Drug D**.~~



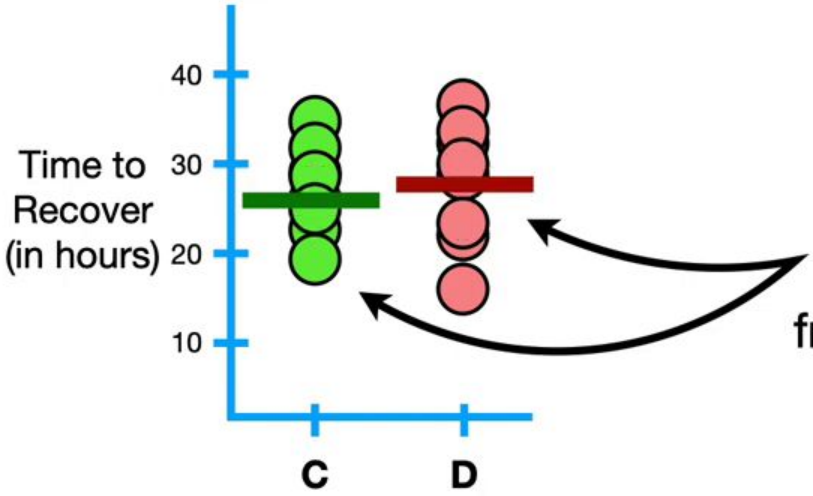
Alternatively, we learned that if little random things could easy shift the result from being in favor one drug to another…

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



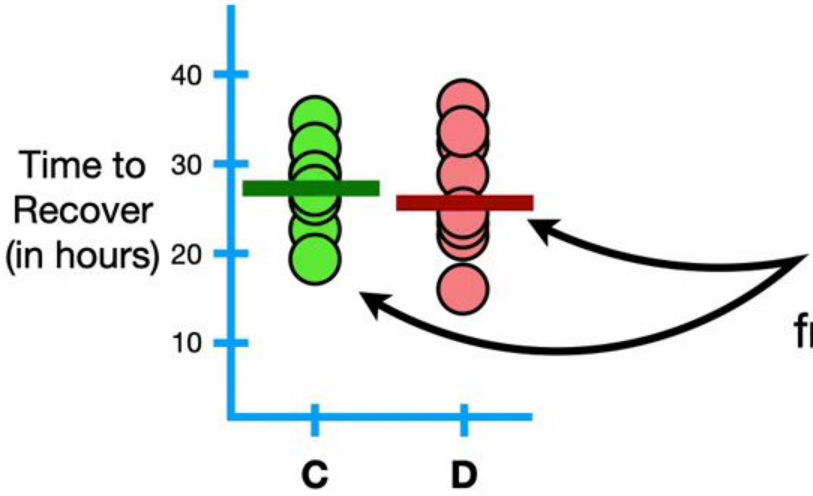
Alternatively, we learned that if little random things could easy shift the result from being in favor one drug to another…

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



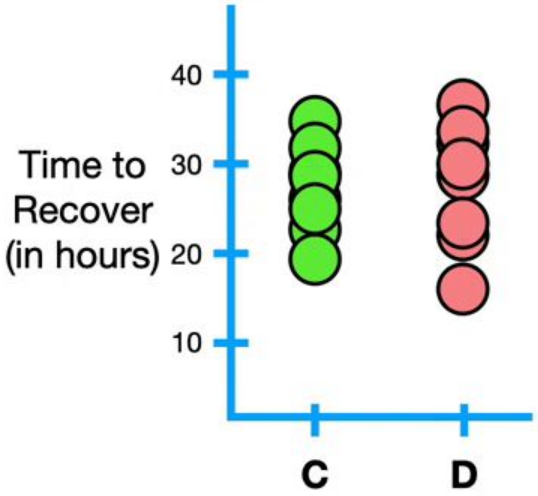
Alternatively, we learned that if little random things could easy shift the result from being in favor one drug to another…

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



Alternatively, we learned that if little random things could easy shift the result from being in favor one drug to another…

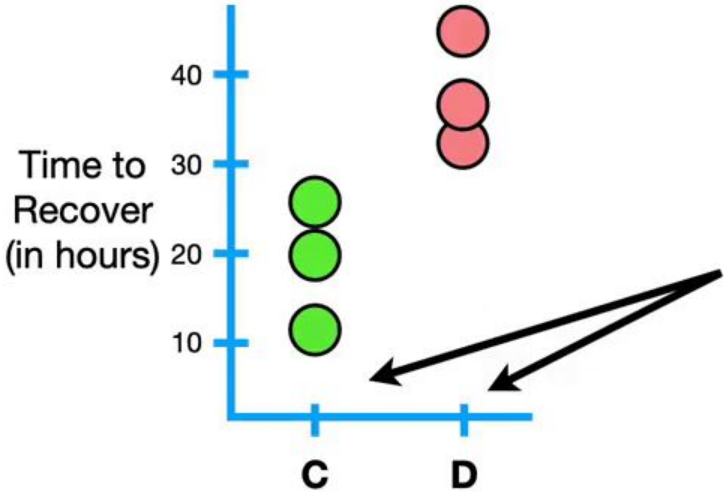
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



…then we would ***fail to reject the Null Hypothesis***.

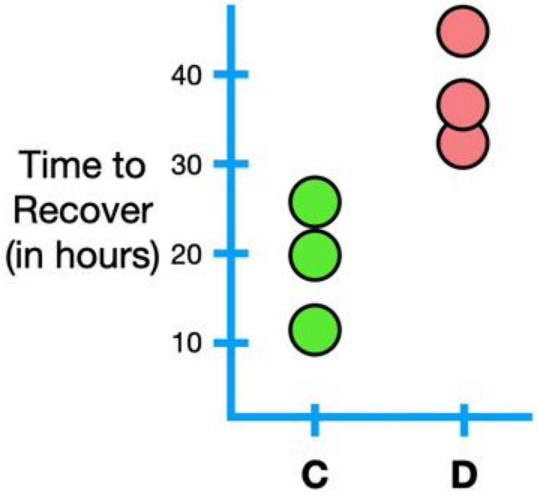
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

Now that we are done with our review, let's talk about the **Alternative Hypothesis**.



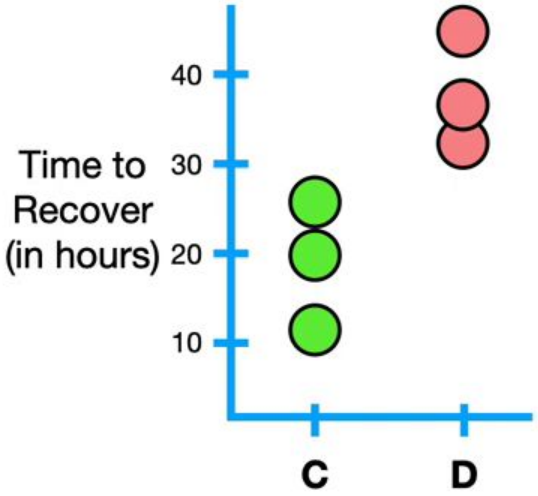
First, here is some data that shows how quickly people taking drugs **C** and **D** recovered from a virus.

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



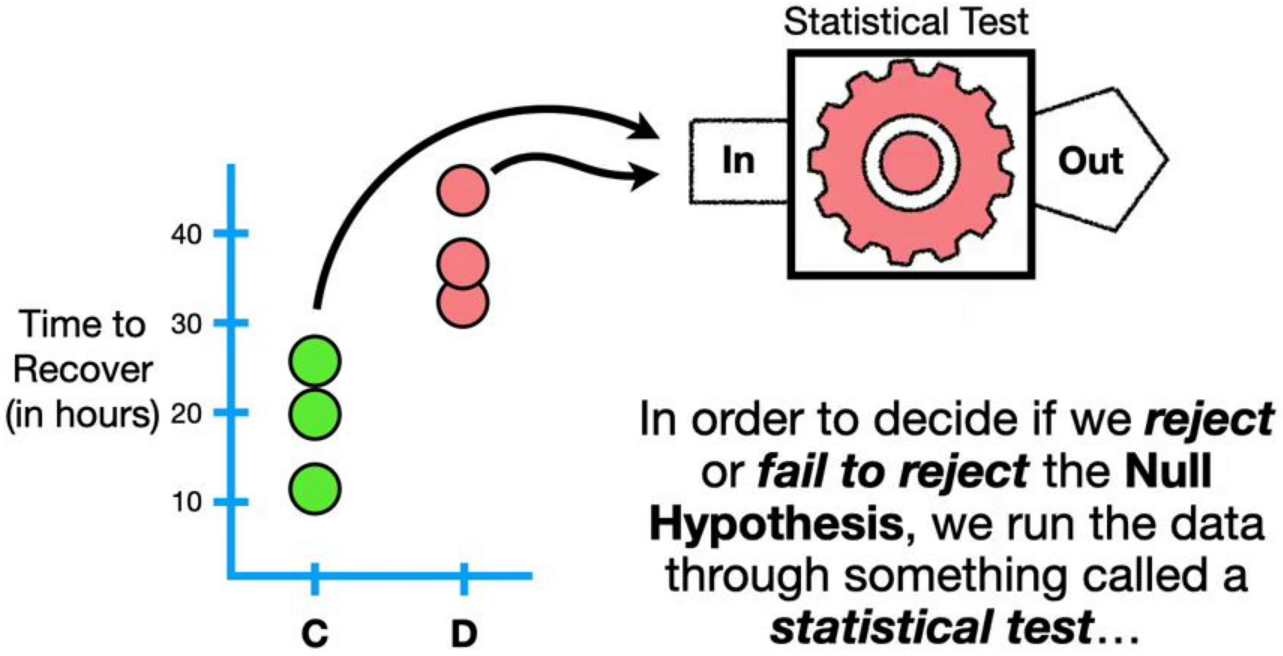
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

The goal of collecting all of this data is to determine if we should ***reject*** or ***fail to reject*** the **Null Hypothesis**.



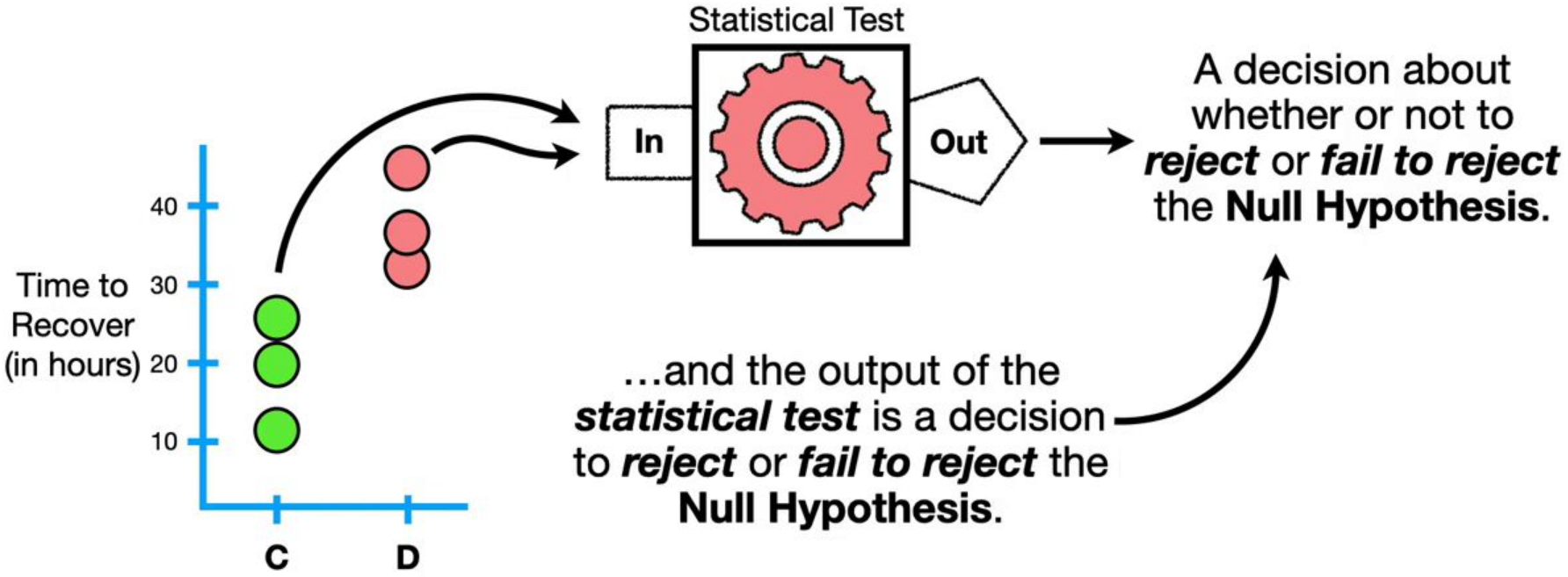
The goal of collecting all of this data is to determine if we should ***reject*** or ***fail to reject*** the **Null Hypothesis**.

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

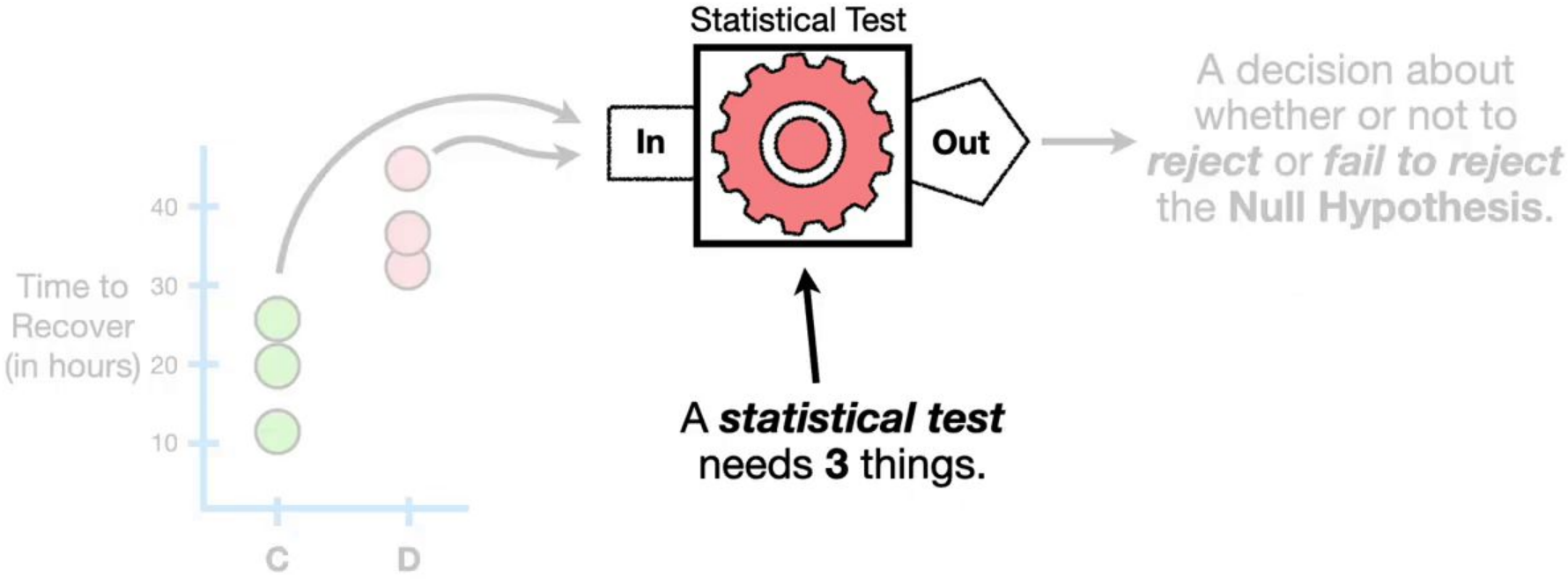


In order to decide if we ***reject*** or ***fail to reject*** the **Null Hypothesis**, we run the data through something called a ***statistical test***…

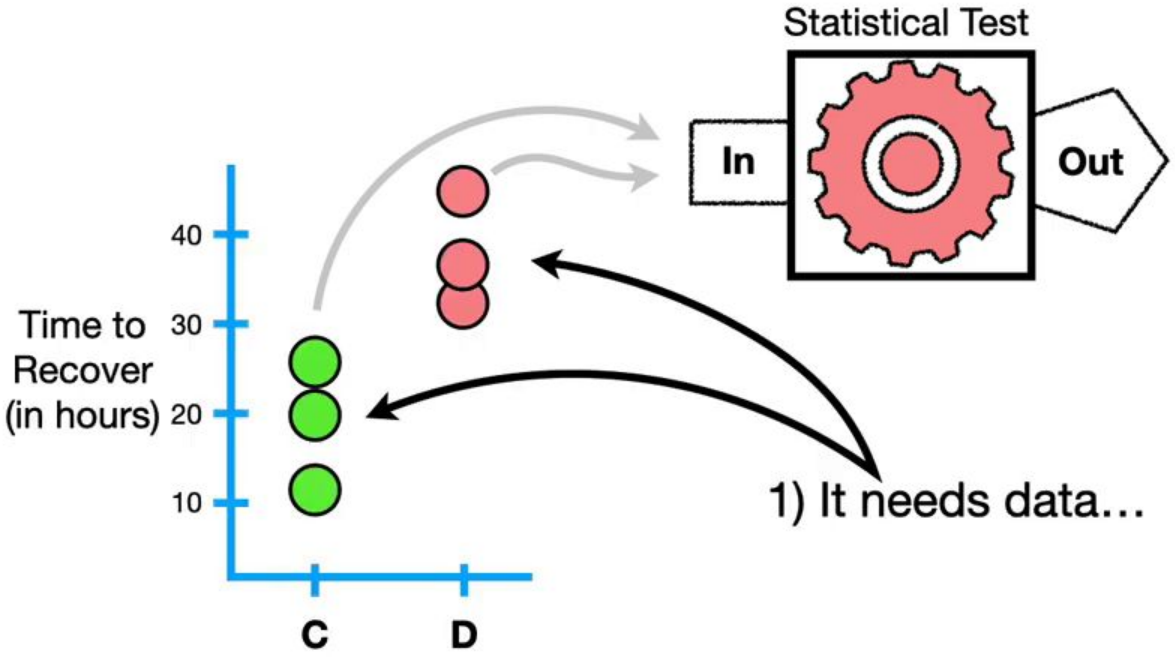
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

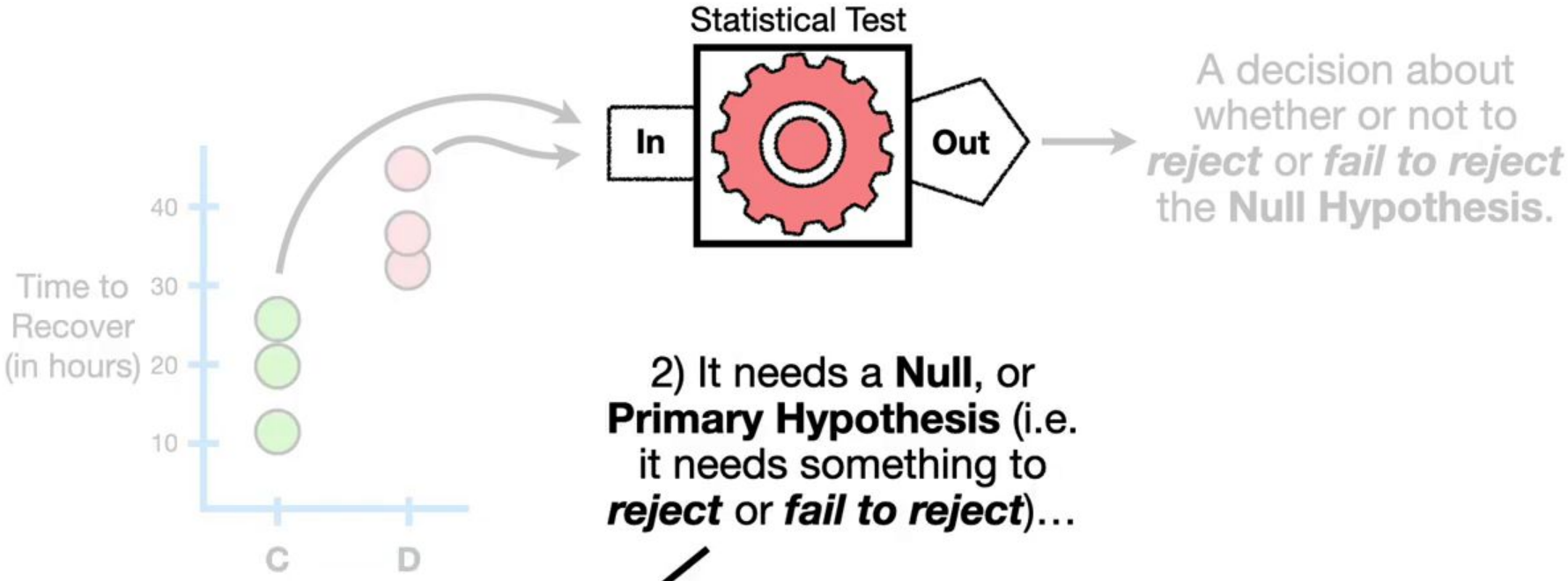


There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



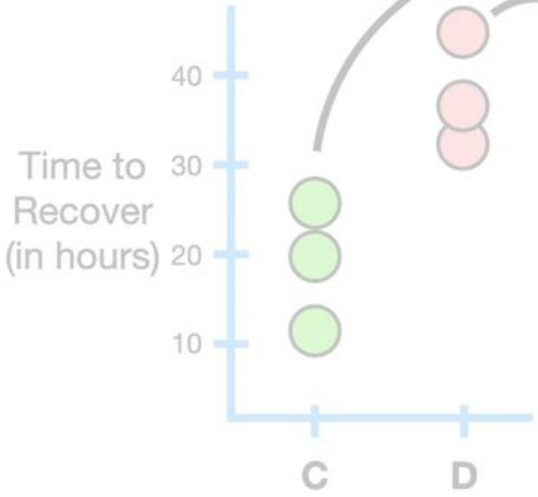
A decision about whether or not to ***reject*** or ***fail to reject*** the **Null Hypothesis**.

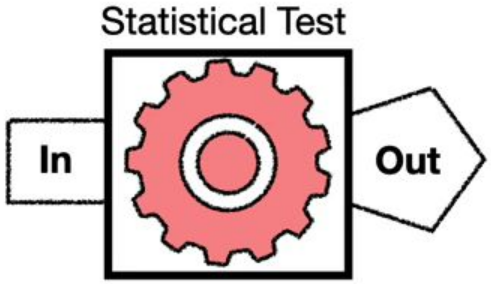
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



2) It needs a **Null**, or **Primary Hypothesis** (i.e. it needs something to ***reject*** or ***fail to reject***)…

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.





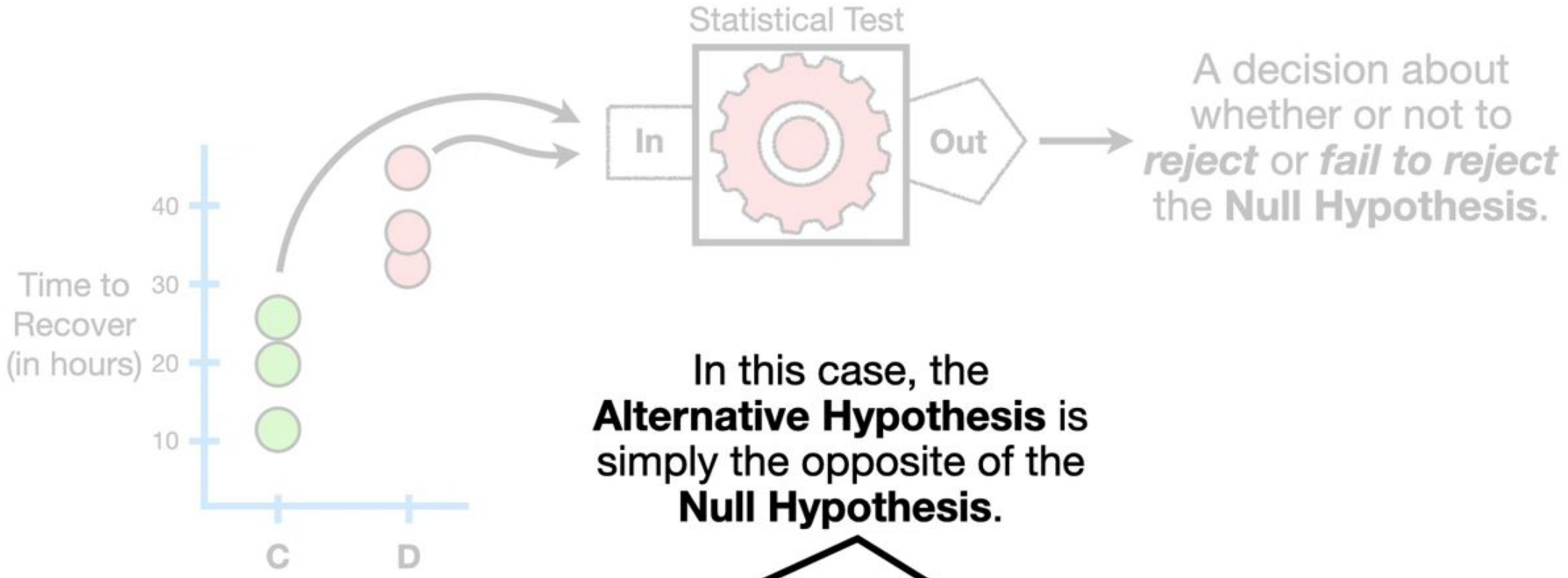
A decision about whether or not to ***reject*** or ***fail to reject*** the **Null Hypothesis**.

…and 3) It needs an **Alternative Hypothesis**.

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.

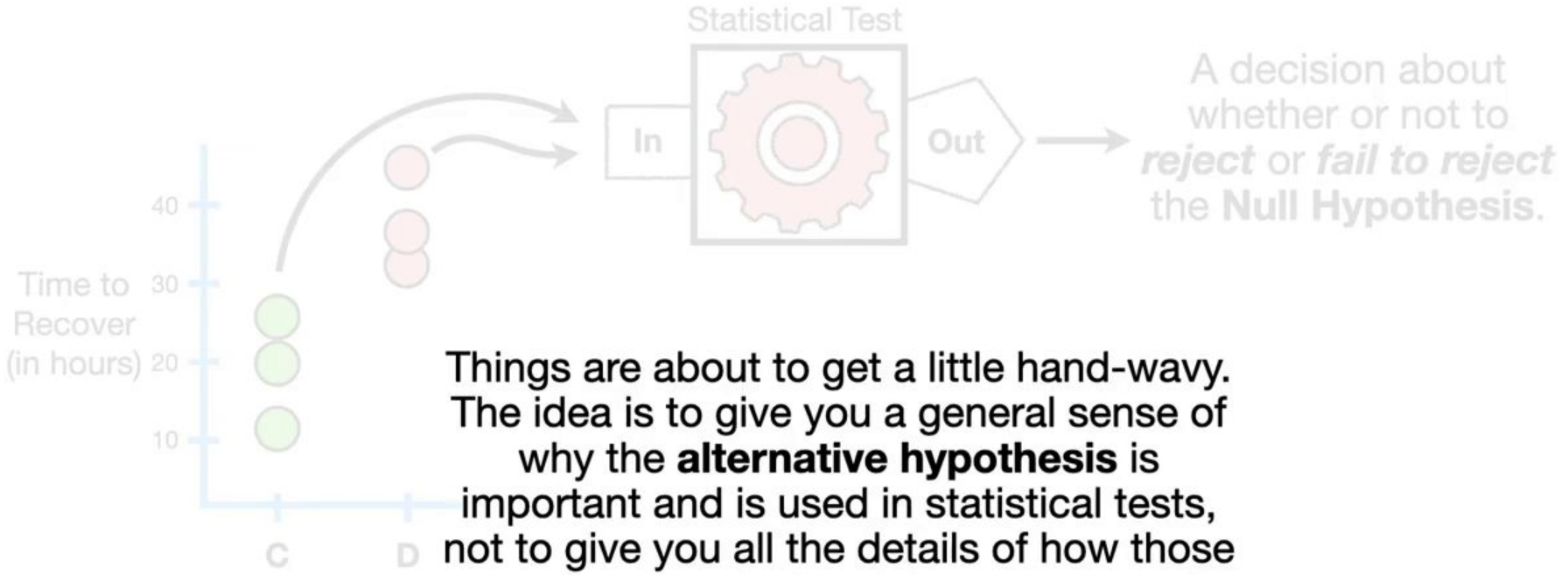


In this case, the **Alternative Hypothesis** is simply the opposite of the **Null Hypothesis**.

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.

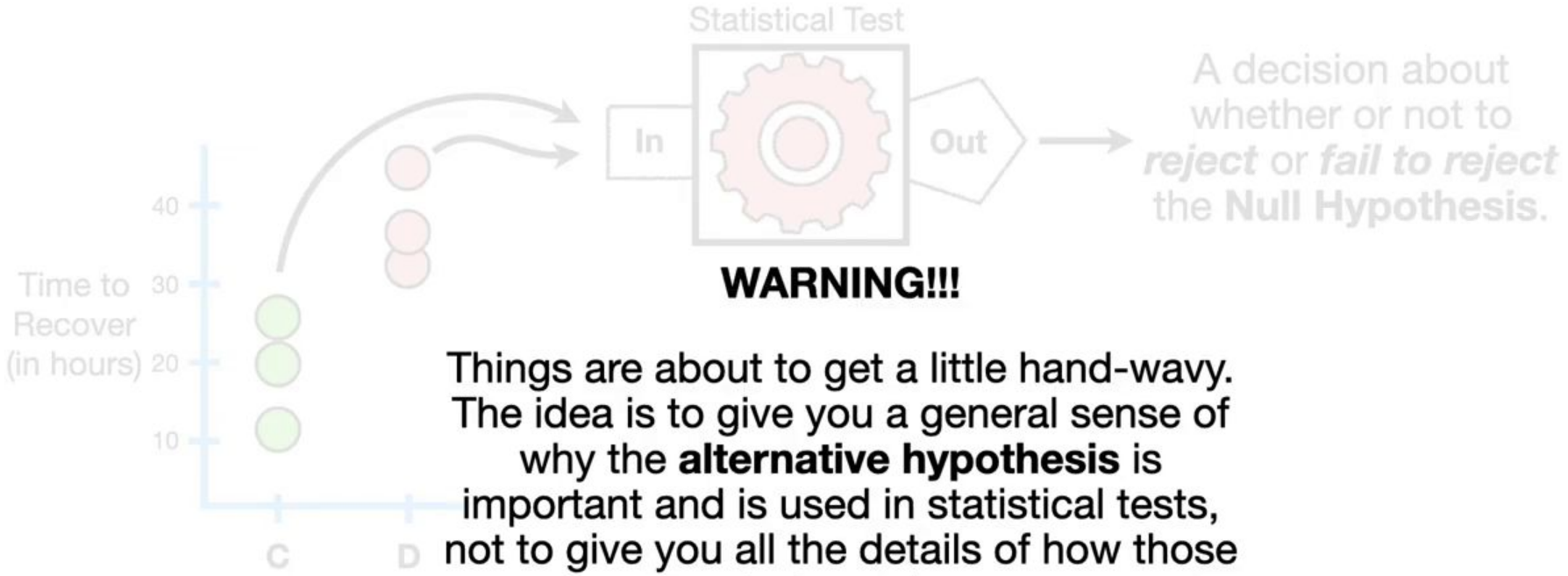


Things are about to get a little hand-wavy. The idea is to give you a general sense of why the **alternative hypothesis** is important and is used in statistical tests, not to give you all the details of how those tests work.

Alternative Hypothesis

There *is a difference* in recovery times between **Drug C** and **Drug D**.

There is *no difference* in recovery times between **Drug C** and **Drug D**.



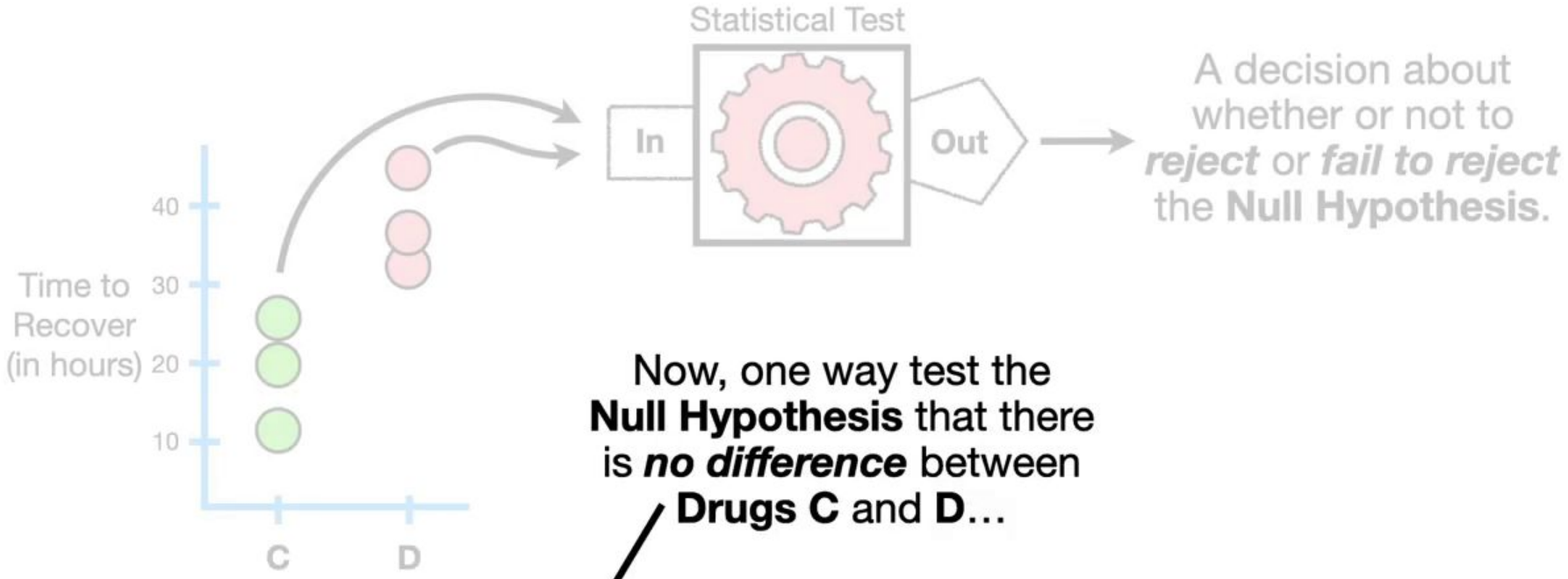
WARNING!!!

Things are about to get a little hand-wavy. The idea is to give you a general sense of why the **alternative hypothesis** is important and is used in statistical tests, not to give you all the details of how those tests work.

Alternative Hypothesis

There *is a difference* in recovery times between **Drug C** and **Drug D**.

There is *no difference* in recovery times between **Drug C** and **Drug D**.

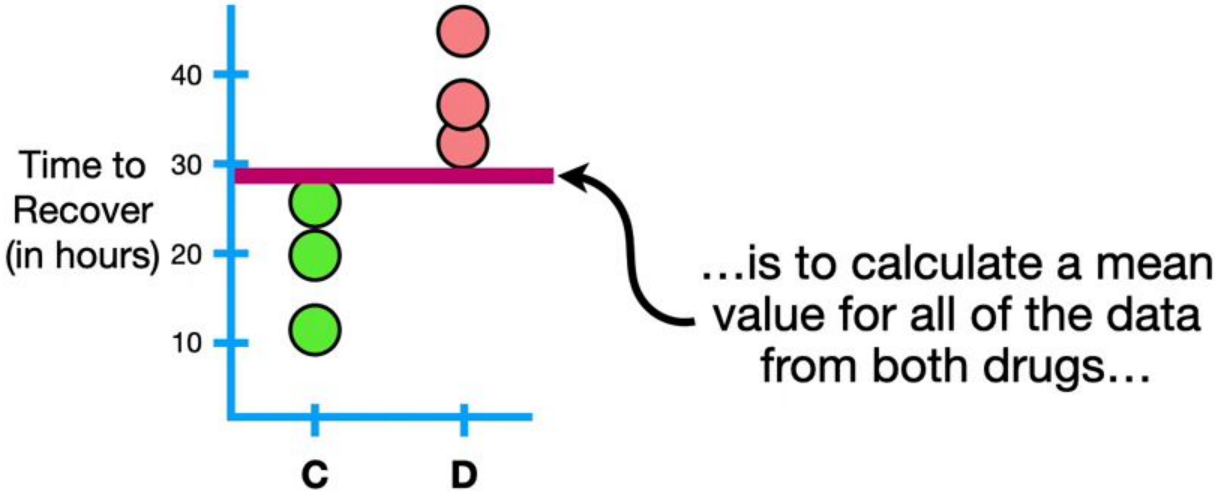


Now, one way test the **Null Hypothesis** that there is ***no difference*** between **Drugs C** and **D**…

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

Alternative Hypothesis

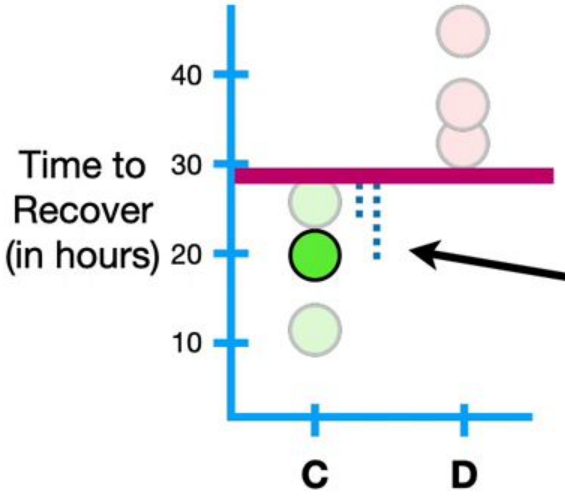
There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.

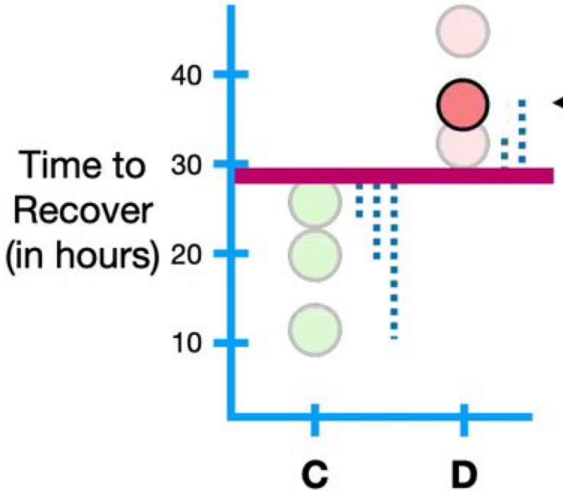


…and calculate the distances between each observation and the mean…

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.

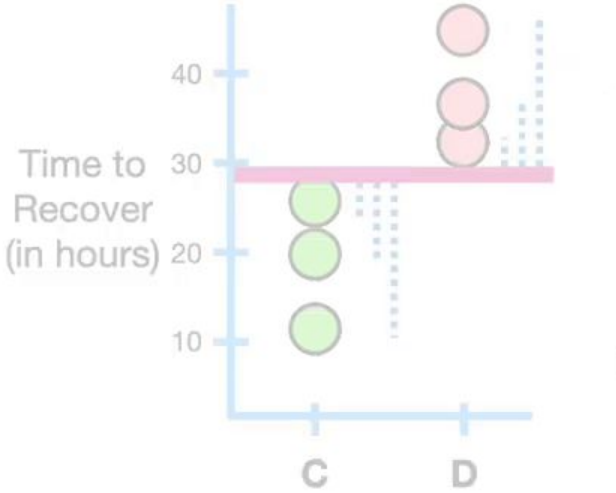


…and calculate the distances between each observation and the mean…

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

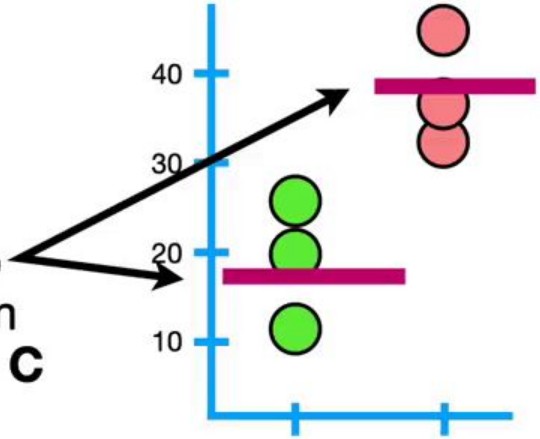
Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



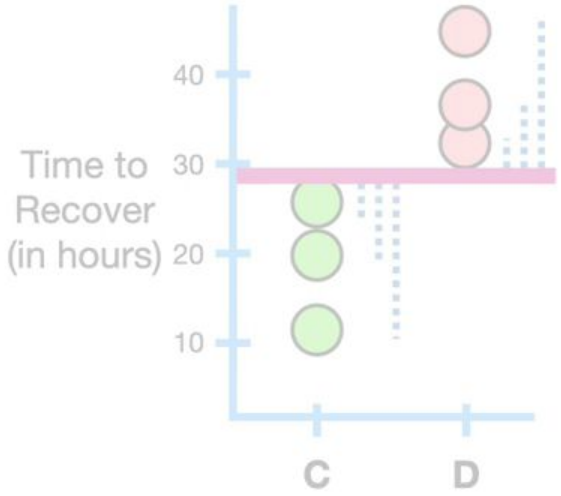
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

…and compare those to distances calculated from individual means for **Drug C** and **Drug D**.

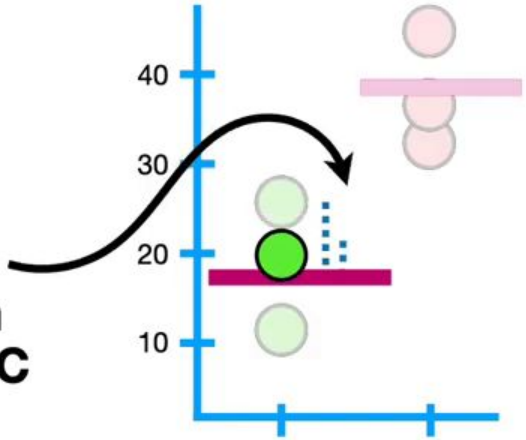


Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



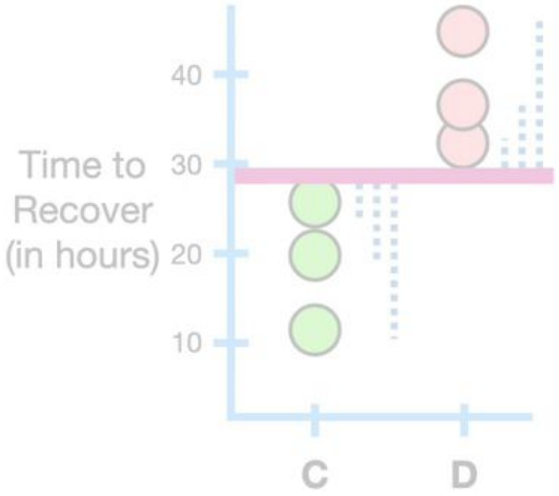
…and compare those to distances calculated from individual means for **Drug C** and **Drug D**.



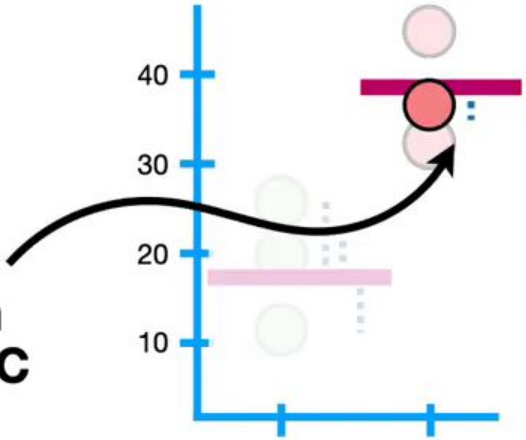
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



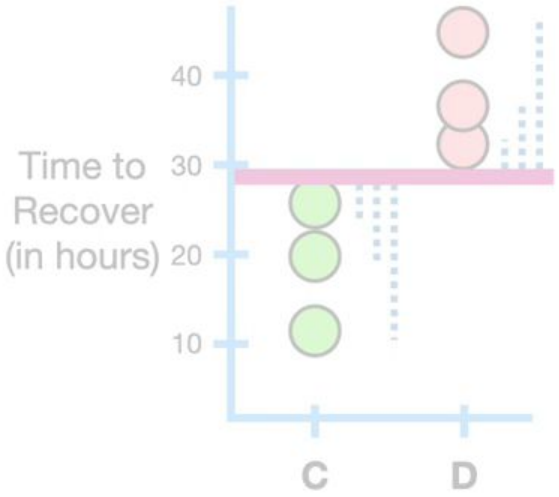
…and compare those to distances calculated from individual means for **Drug C** and **Drug D**.



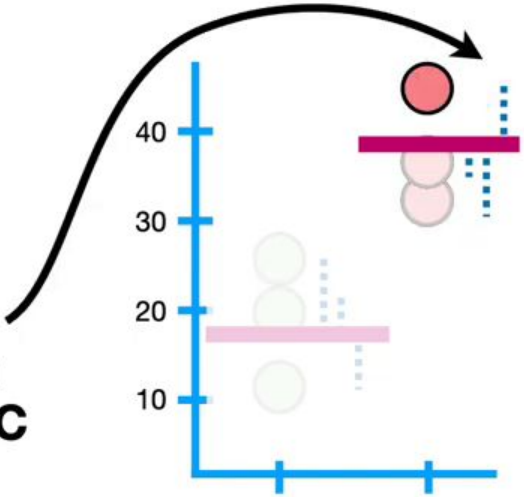
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



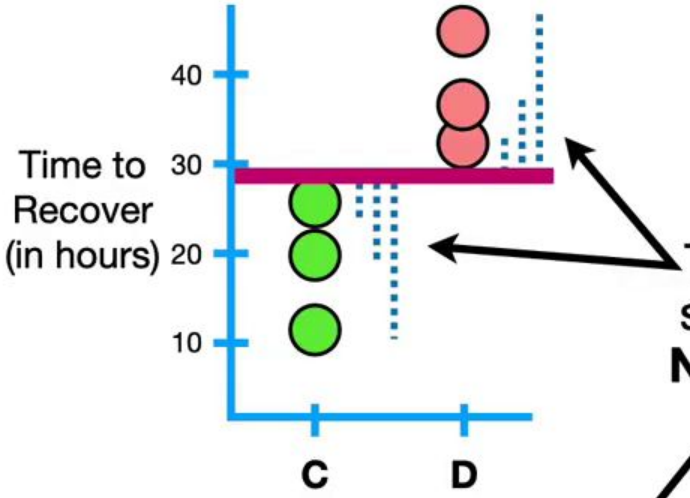
…and compare those to distances calculated from individual means for **Drug C** and **Drug D**.



There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

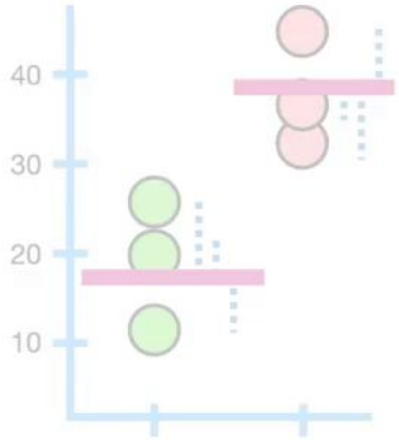
Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



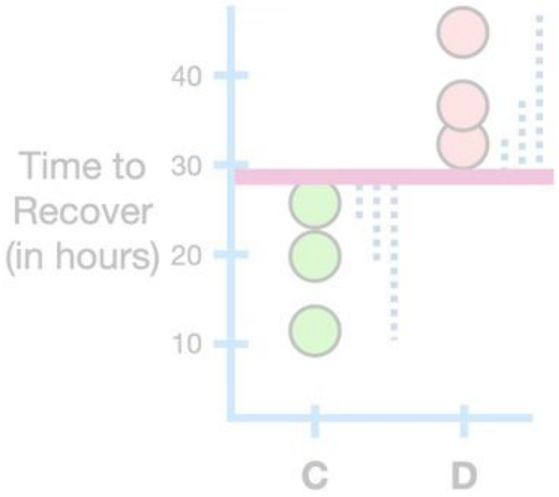
The distances around the single mean represent the **Null Hypothesis** that there is ***no difference***…

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



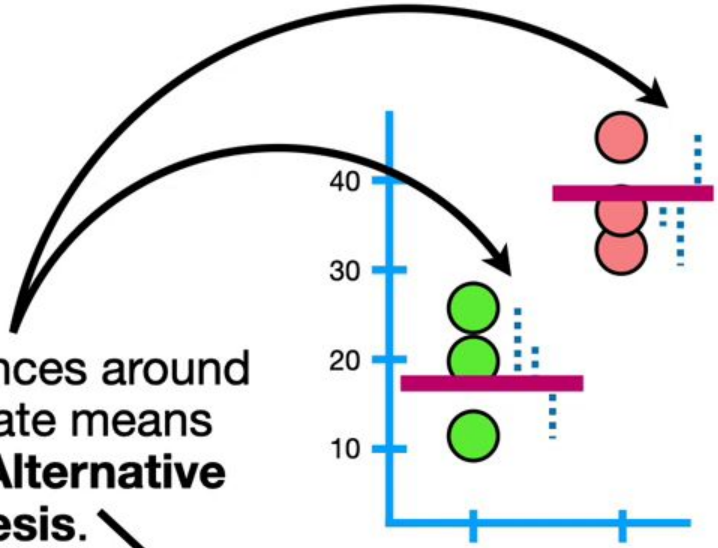
Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



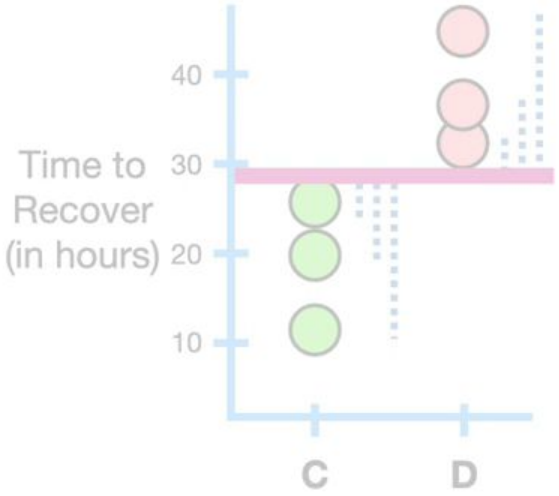
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

…and the distances around the two separate means represent the **Alternative Hypothesis**.



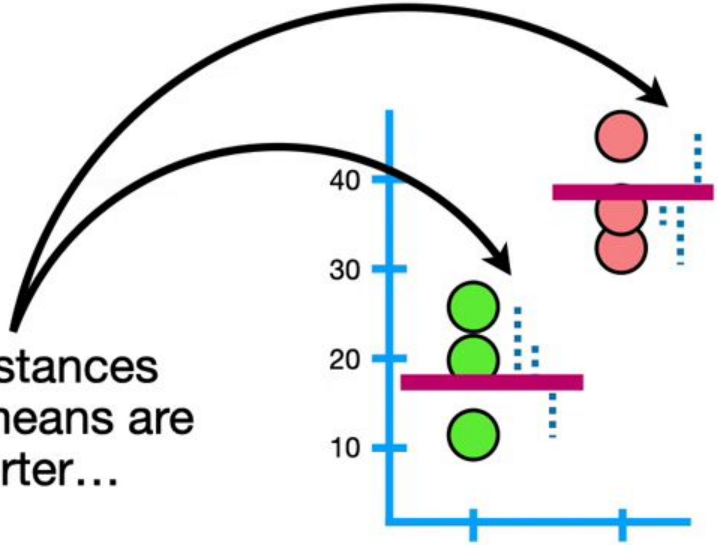
Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



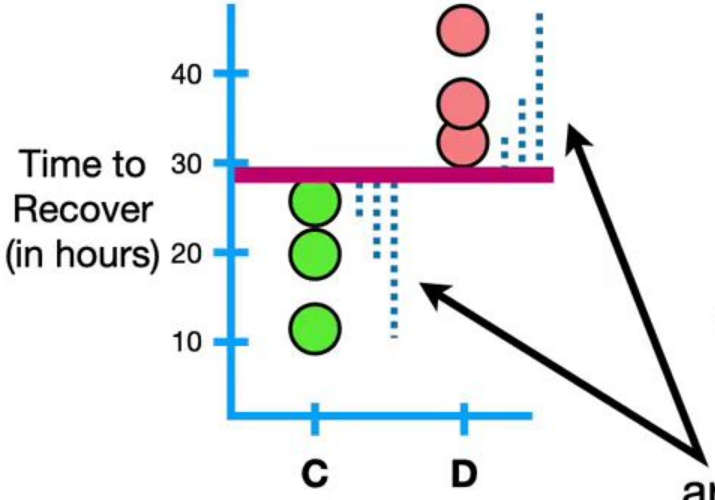
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

If the the distances around two means are much shorter…



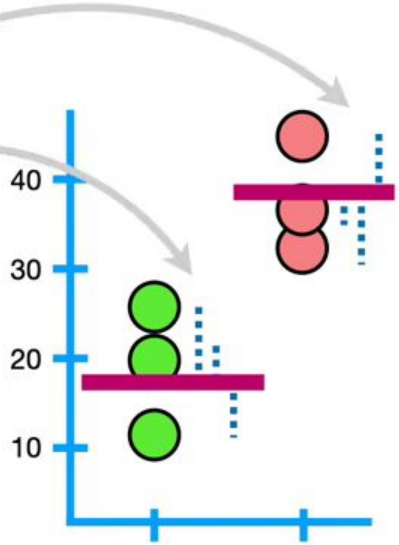
Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



If the the distances around two means are much shorter…

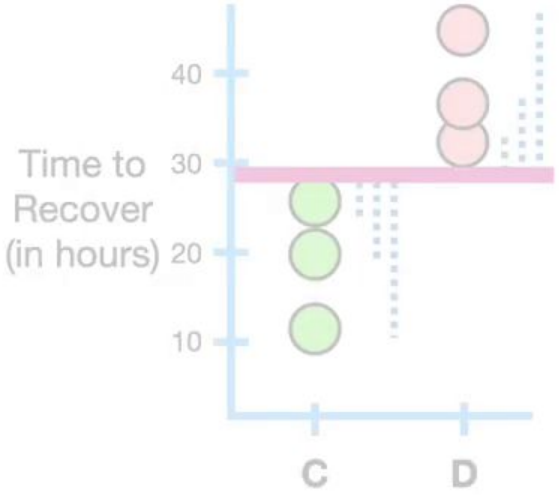
…than the distances around the single mean…



Alternative Hypothesis

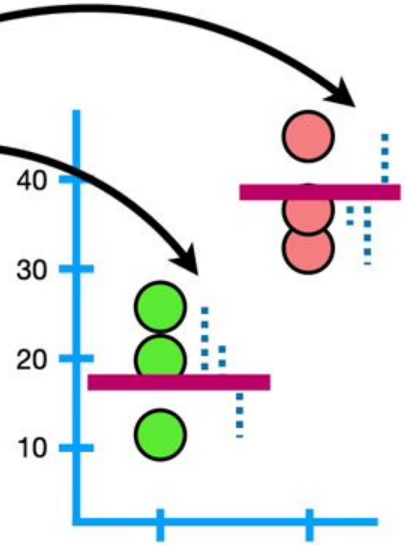
There ***is a difference*** in recovery times between **Drug C** and **Drug D**.

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



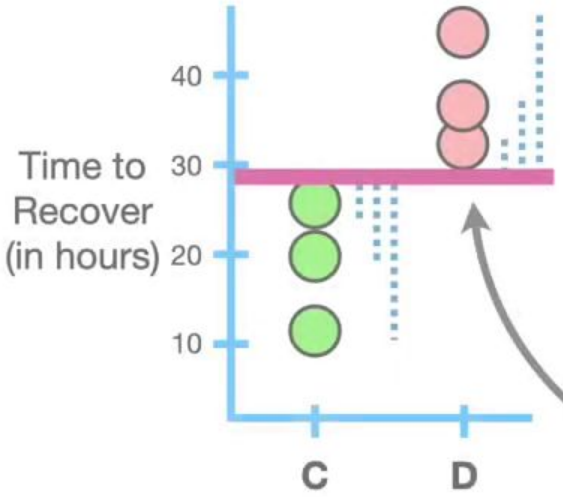
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

...then that suggests that using two means to summarize the data makes more sense than using one.

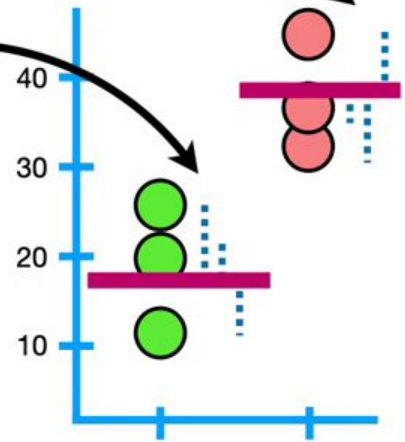


Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



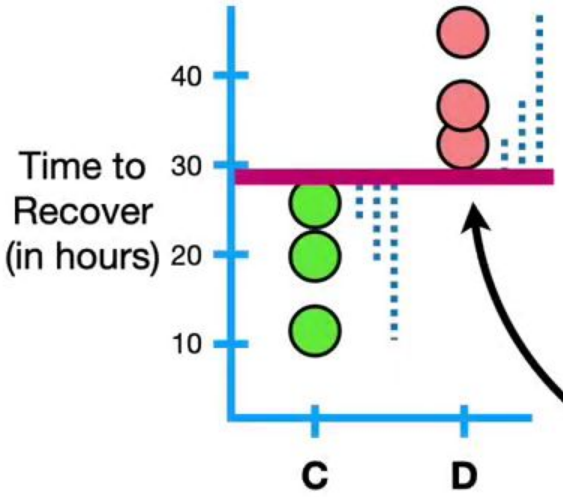
…then that suggests that using two means to summarize the data makes more sense than using one.



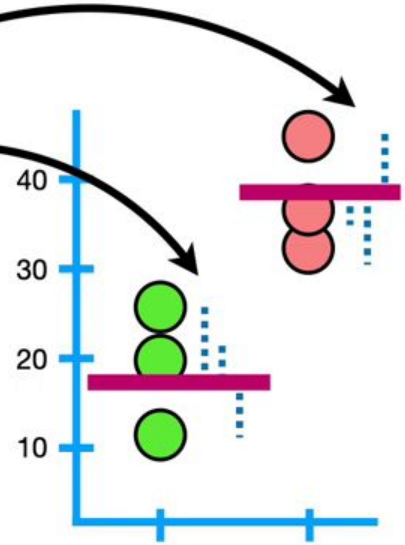
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



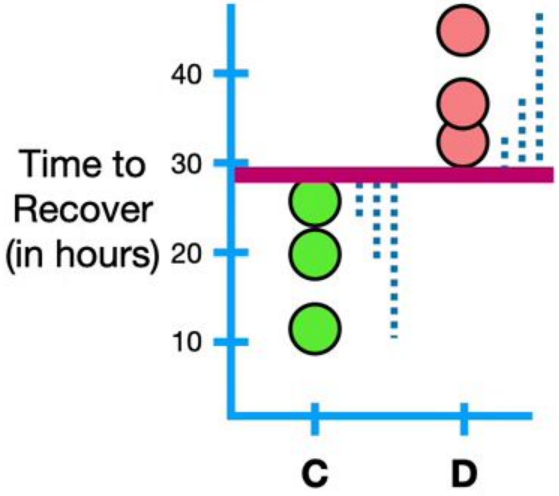
…then that suggests that using two means to summarize the data makes more sense than using one.



Alternative Hypothesis

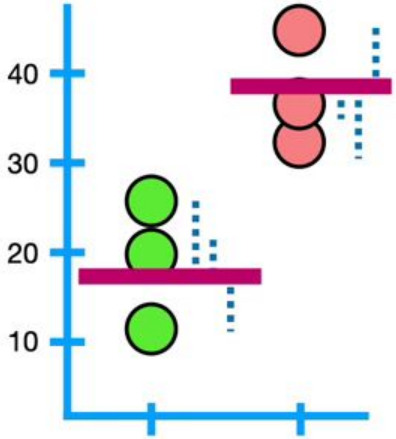
There ***is a difference*** in recovery times between **Drug C** and **Drug D**.

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



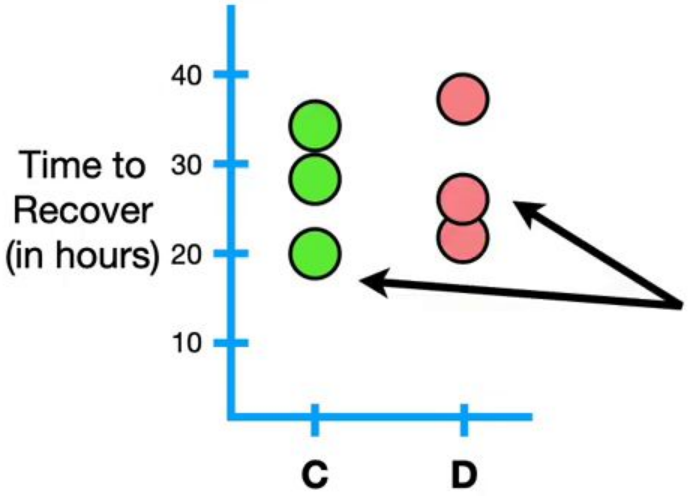
…so we will ***reject*** the **Null Hypothesis.**

~~There is ***no difference*** in recovery times between **Drug C** and **Drug D**.~~



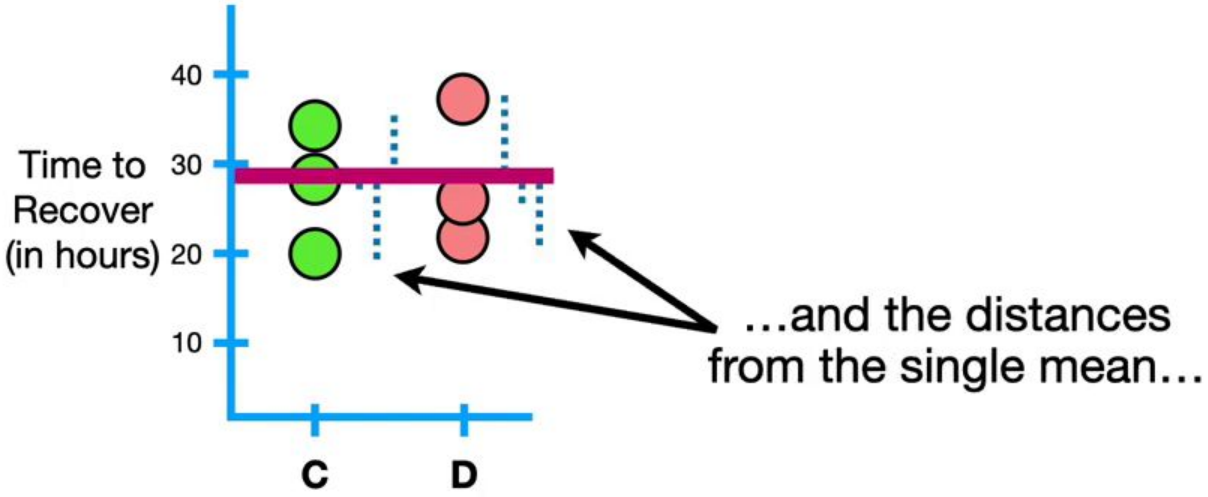
Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.

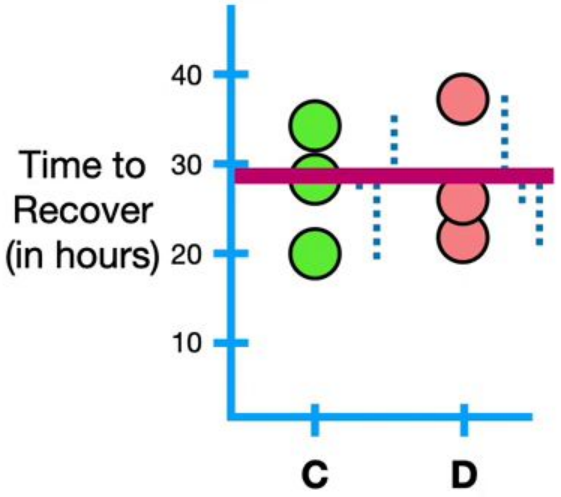


Alternatively, if the data looked like this…

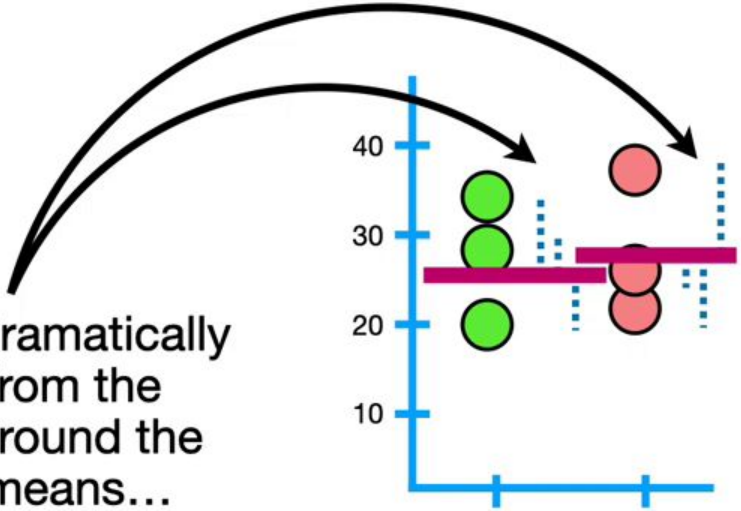
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



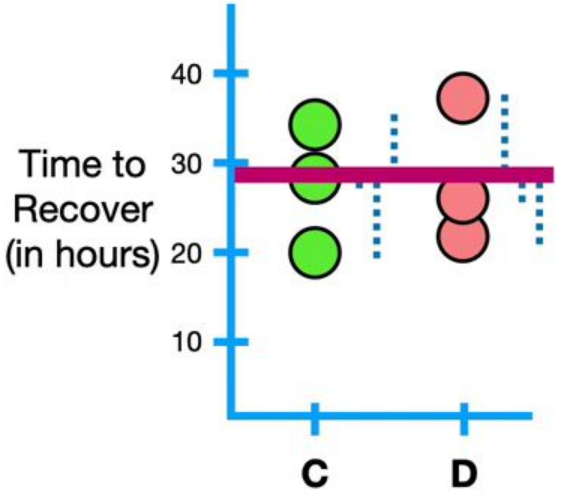
…were not dramatically different from the distances around the separate means…



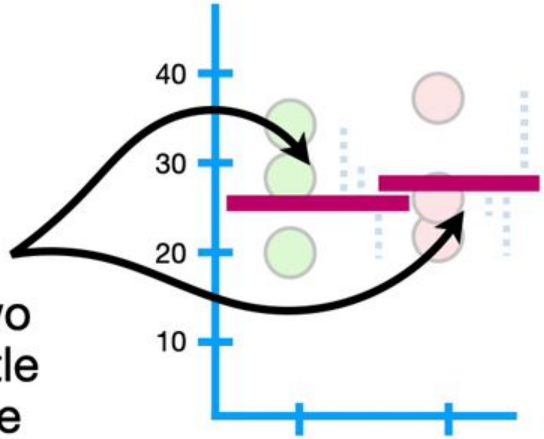
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



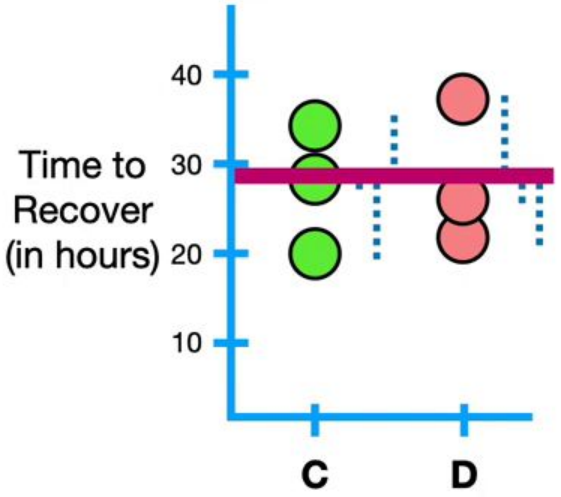
…then that would suggest that the difference between two means only reflects little random things that we can't account for.



There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

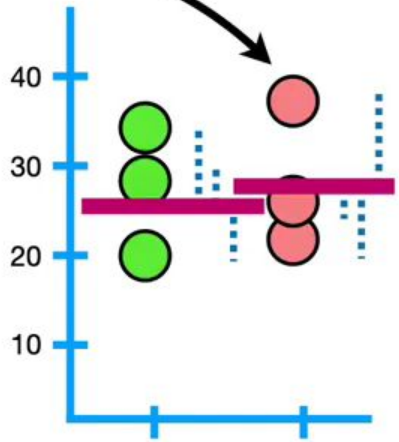
Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



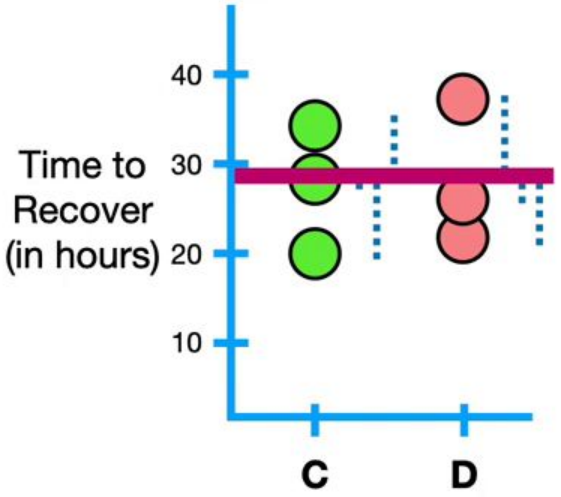
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

For example, it could be that the subtle difference in the means is due to this one guy getting less exercise than everyone else.



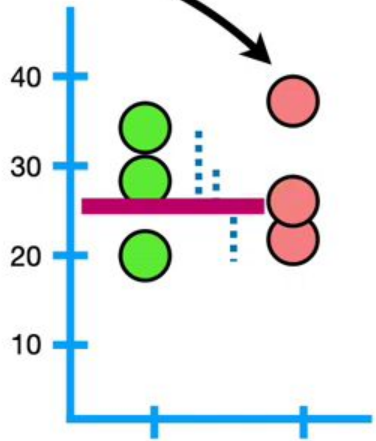
Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



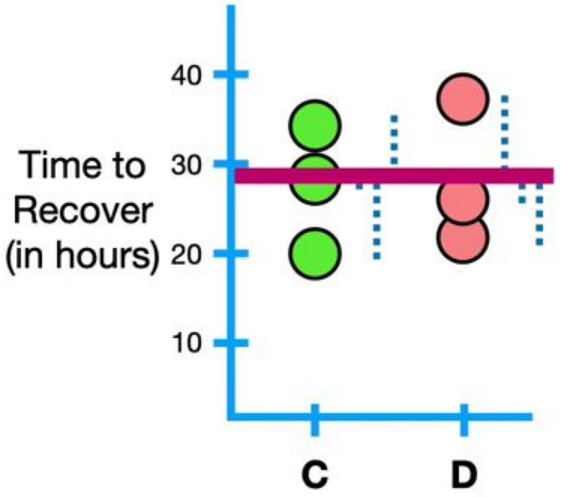
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

If he had exercised just a little bit more, he might have recovered from the illness a little faster…

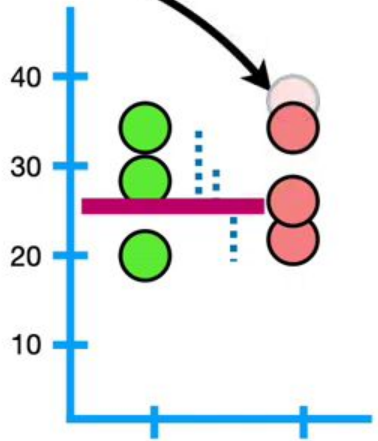


Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



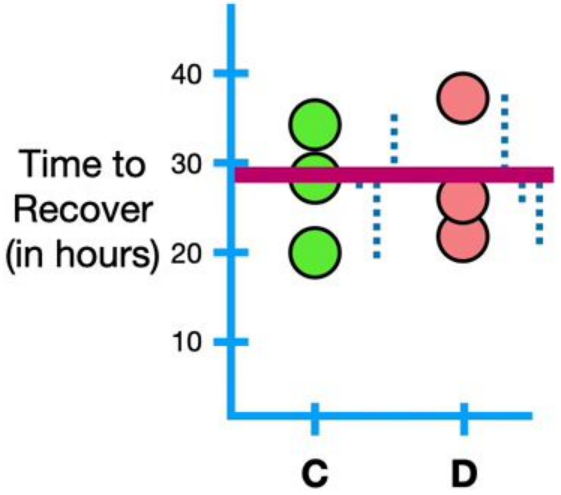
If he had exercised just a little bit more, he might have recovered from the illness a little faster…



There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

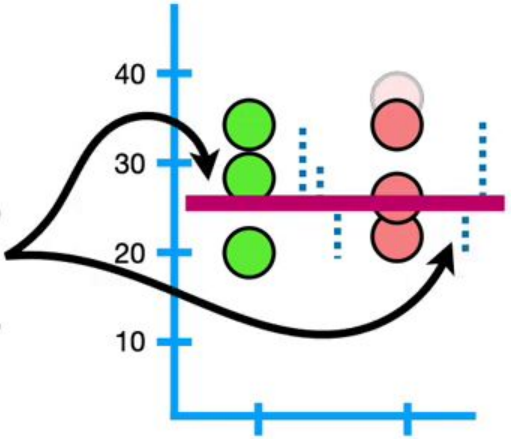
Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



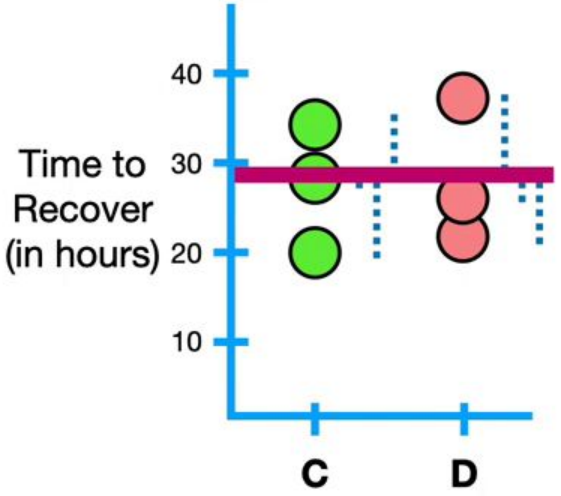
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

…and then we would no longer see a difference between the two means.



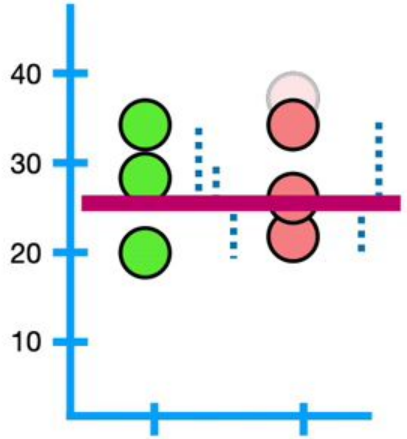
Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



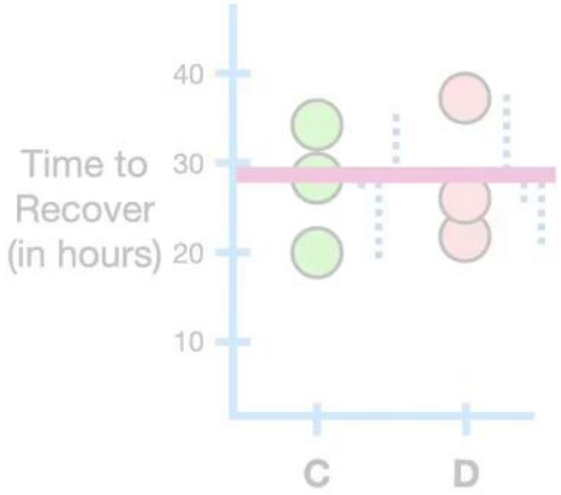
So in this case, we would ***fail to reject*** the **Null Hypothesis**.

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

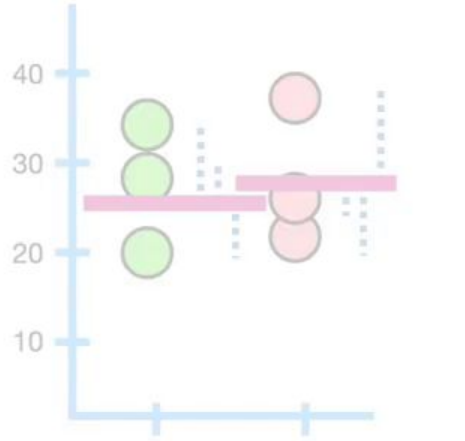


Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



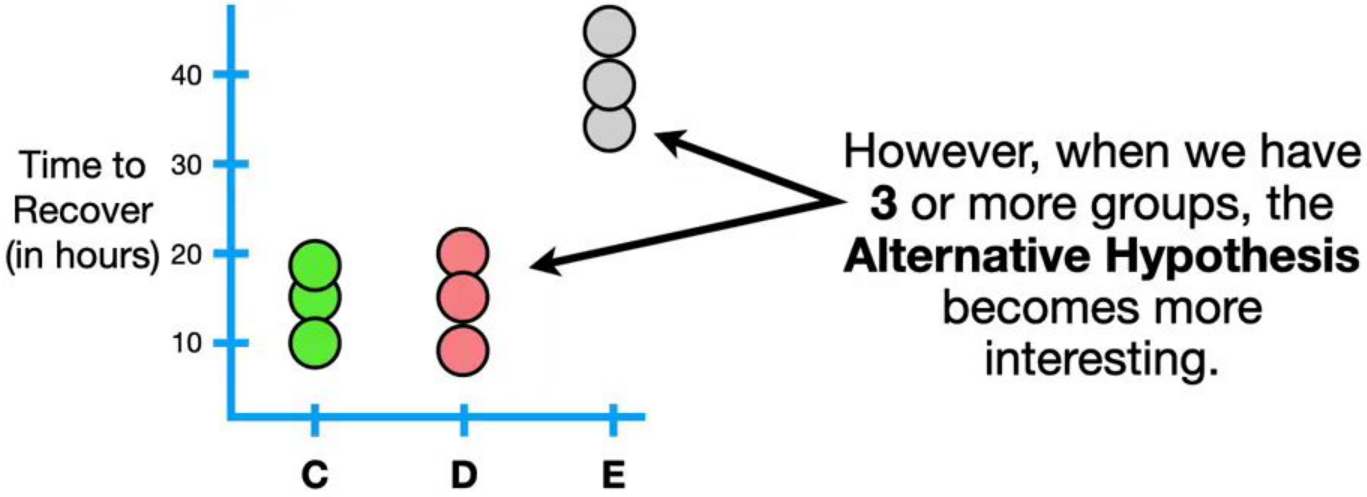
NOTE: When we only have two groups of data, the **Alternative Hypothesis** is pretty obvious because it is simply the opposite of the **Null Hypothesis**.



There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

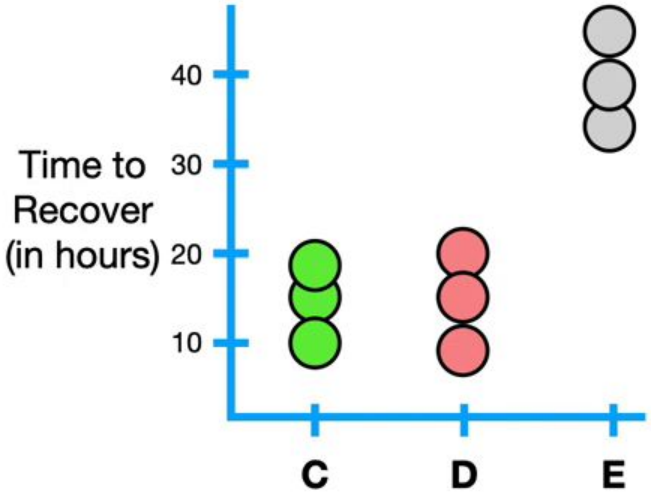
Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



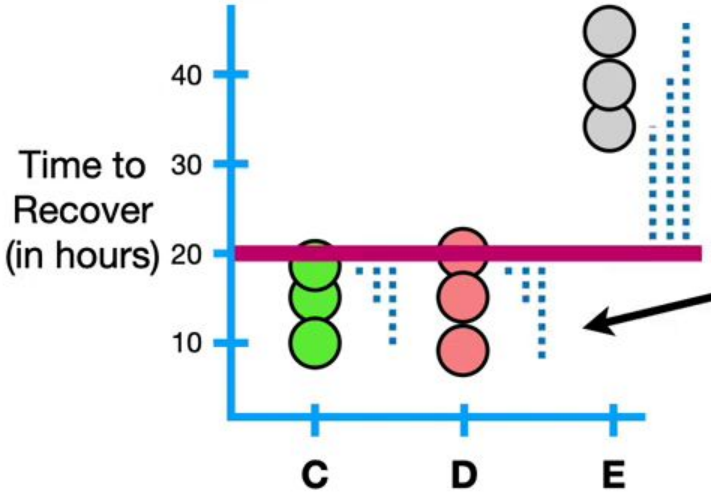
However, when we have **3** or more groups, the **Alternative Hypothesis** becomes more interesting.

There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.



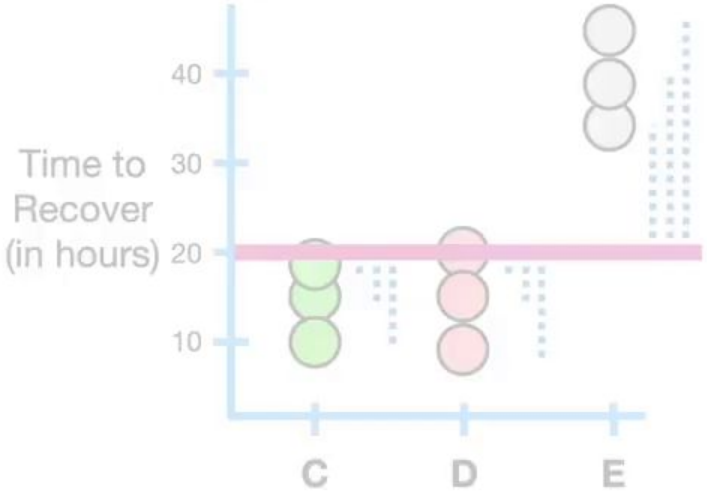
In this case, the **Null Hypothesis** is that there is ***no difference*** between **Drugs C**, **D** and **E**…

There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.



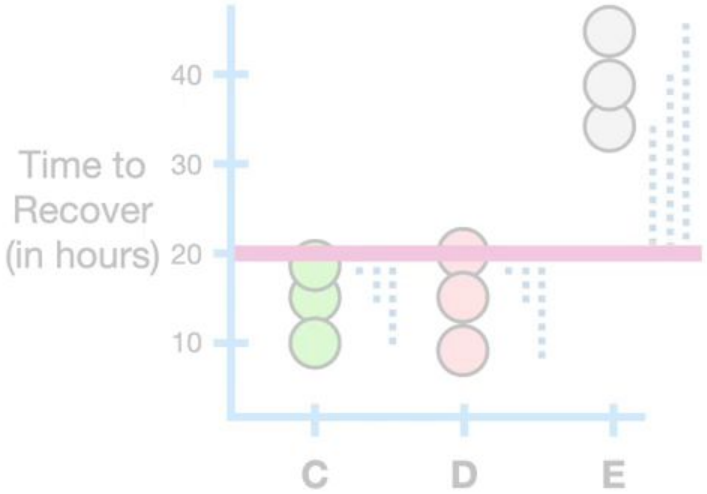
…and like before, we can represent the **Null Hypothesis** by measuring the distances from the data to a single mean value.

There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.



There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.

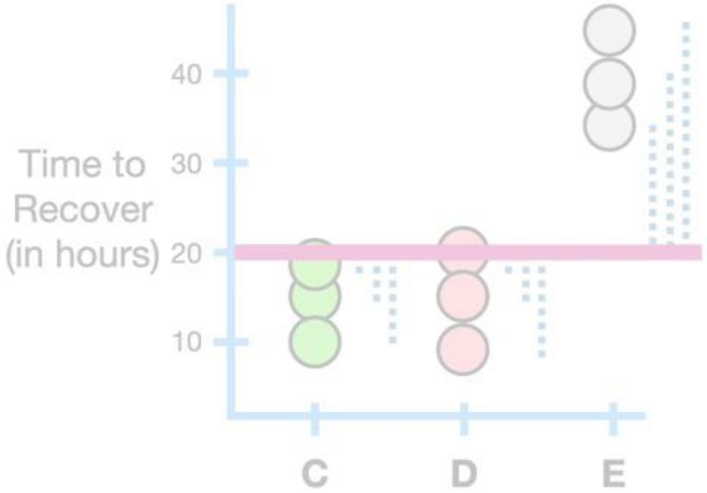
However, now we have choices for the **Alternative Hypothesis**.



There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.

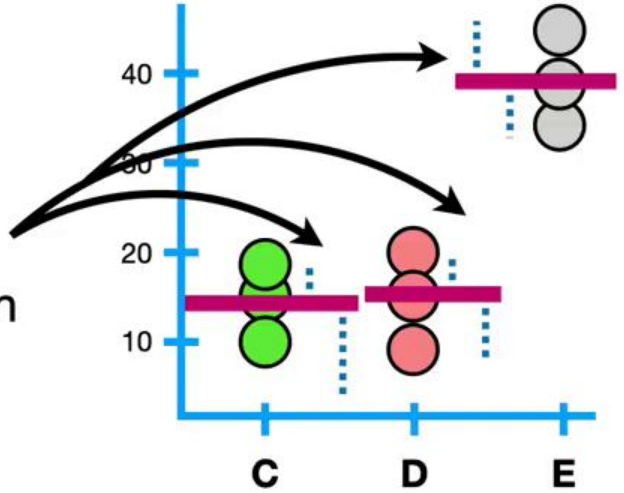
One **Alternative Hypothesis** could be that all three drugs are different…

Drugs C, D and **E** are all different from each other.

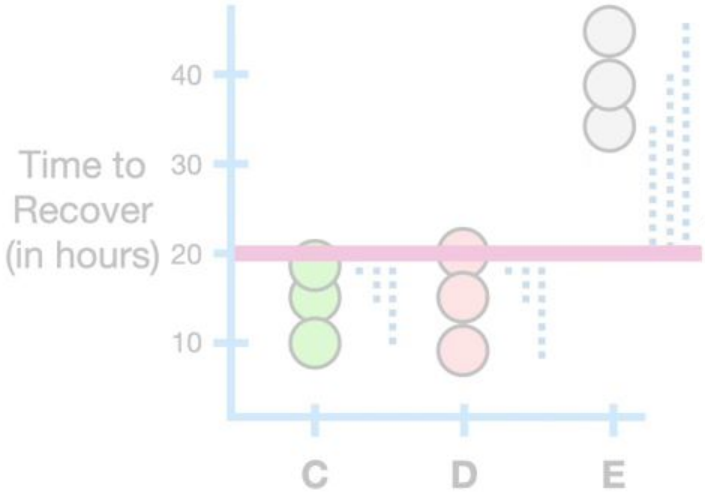


There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.

…and in this case we would measure the distances from a separate mean for each drug.



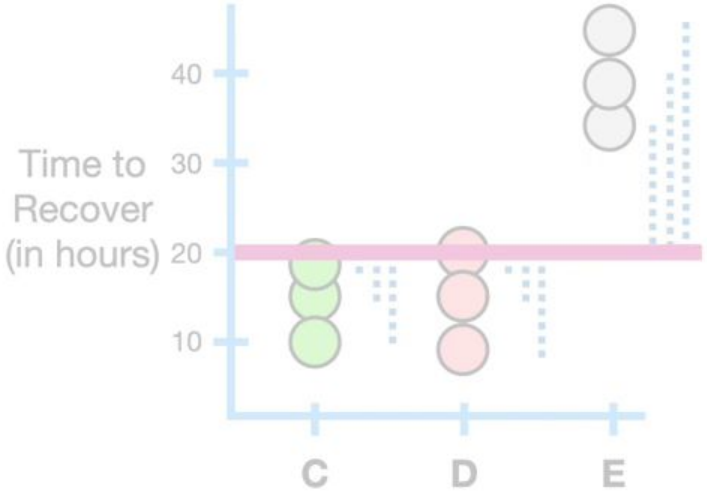
Drugs C, D and **E** are all different from each other.



There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.

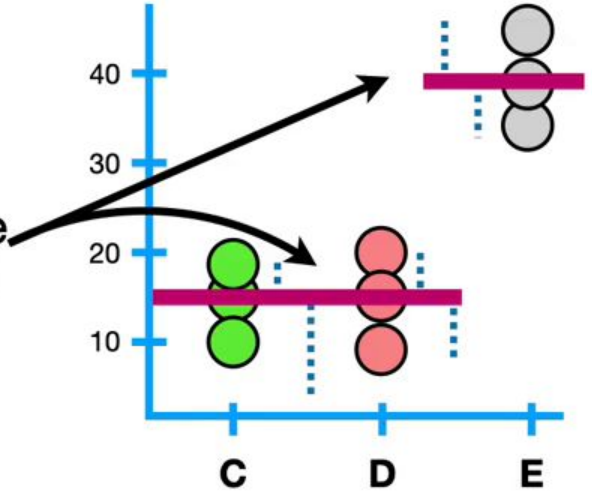
Or the **Alternative Hypothesis** could be that there is no difference between **Drugs C** and **D**, but **Drug E** is ***doing its own thing***.

Drugs C, D are the same and **Drug E** is ***doing its own thing***.

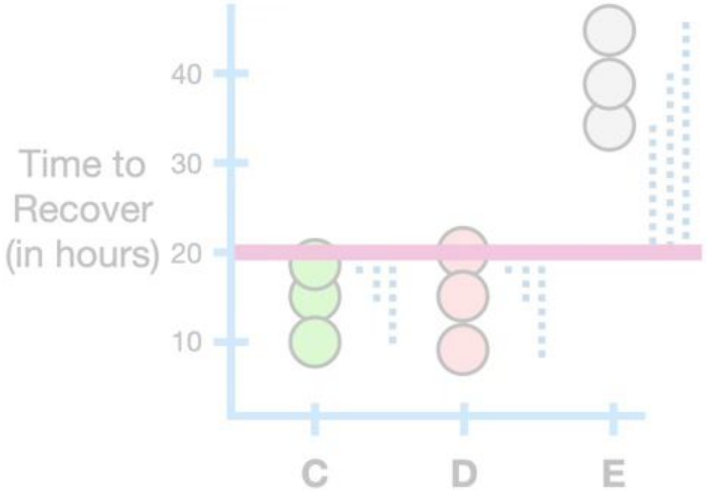


There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.

In this case, we would calculate the distances from a single mean value for **Drugs C** and **D**, and a separate mean for **Drug E**.

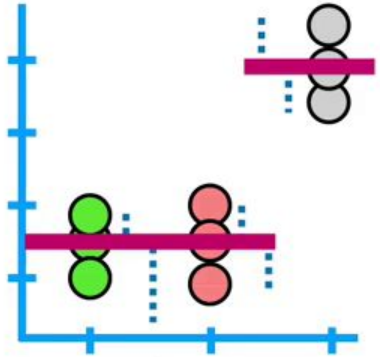


Drugs C, D are the same and **Drug E** is ***doing its own thing***.

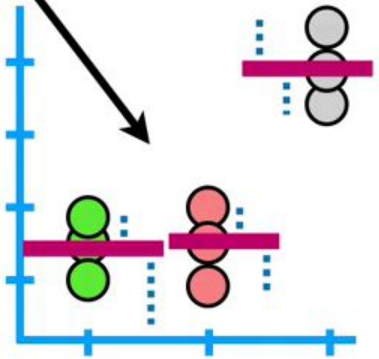


There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.

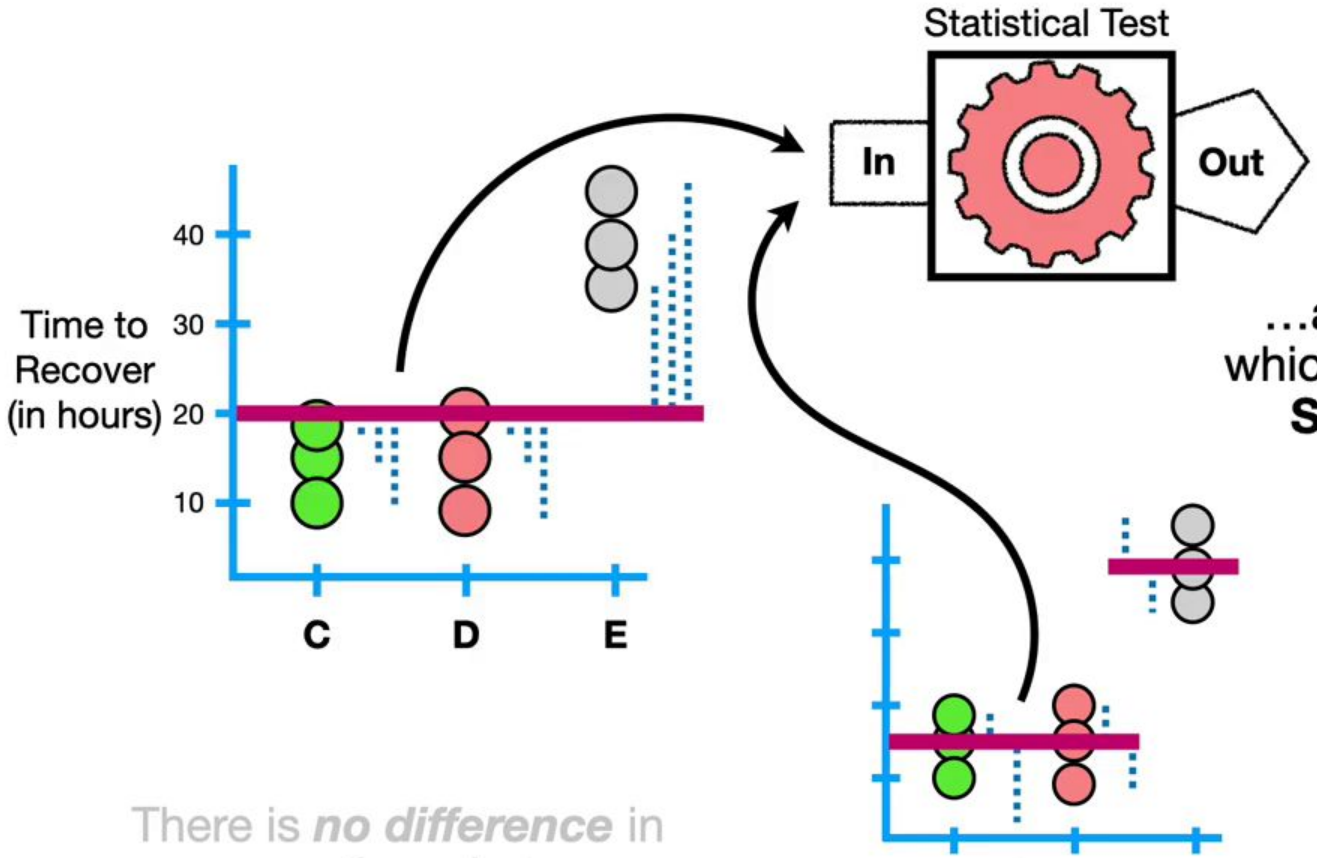
So far we have two different **Alternative Hypotheses**…



Drugs C, D are the same and **Drug E** is ***doing its own thing***.



Drugs C, D and **E** are all different from each other.

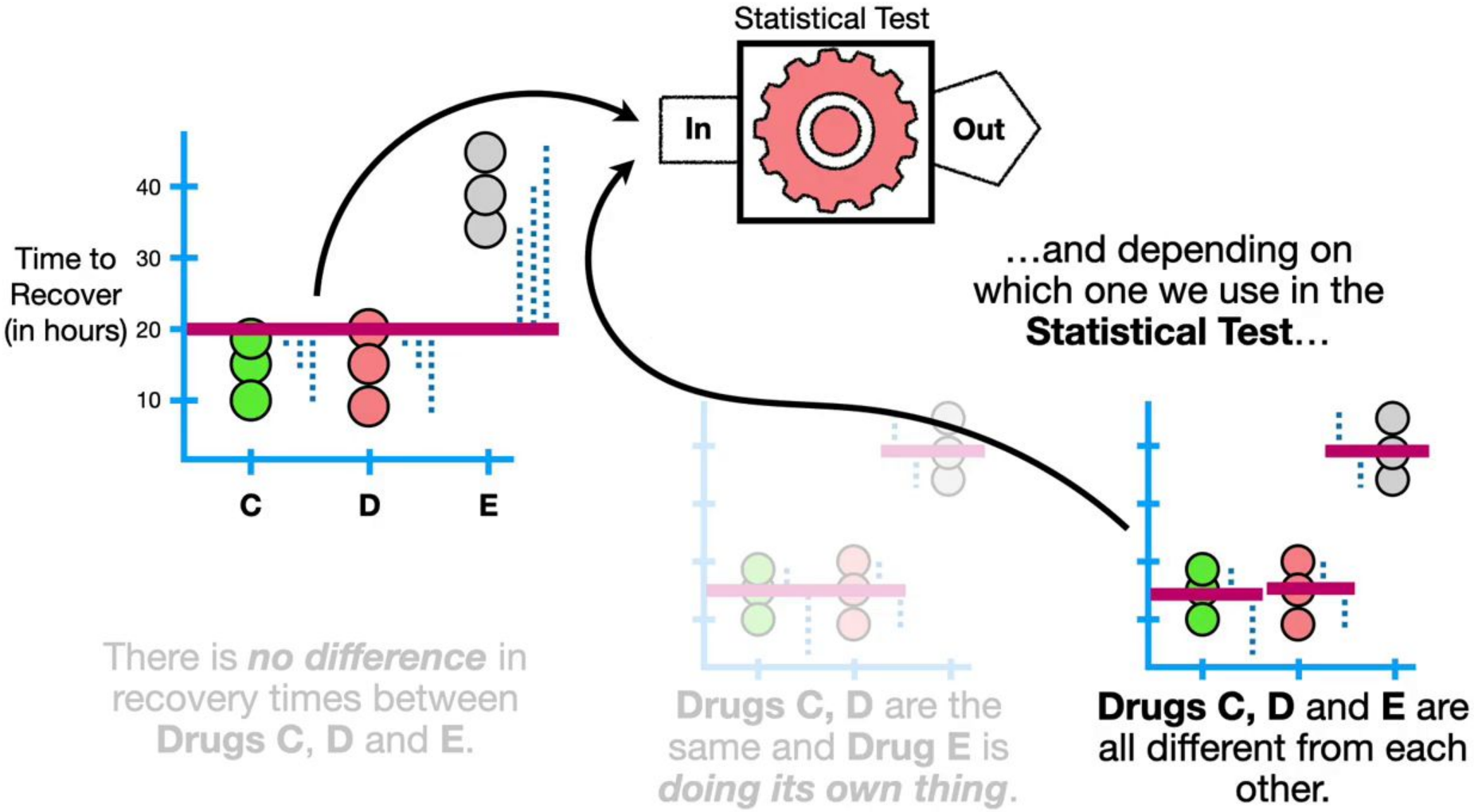


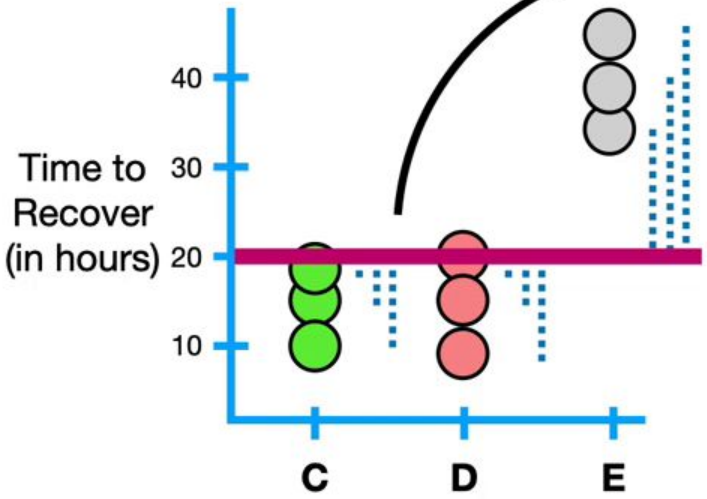
…and depending on which one we use in the **Statistical Test**…

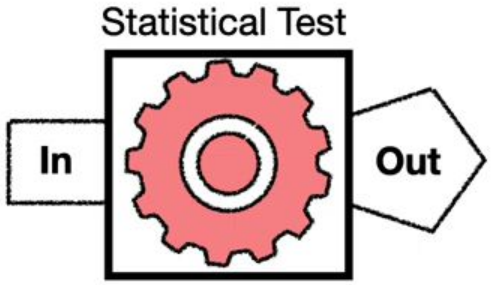
There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.

Drugs C, D are the same and **Drug E** is ***doing its own thing***.

Drugs C, D and **E** are all different from each other.


Statistical Test
In
Out
Time to Recover (in hours)
40
30
20
10
C
D
E
…and depending on which one we use in the **Statistical Test**…
There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.
Drugs C, D are the same and **Drug E** is ***doing its own thing***.
Drugs C, D and **E** are all different from each other.

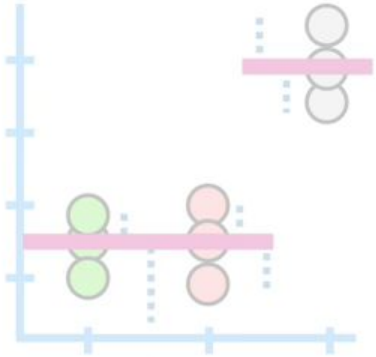




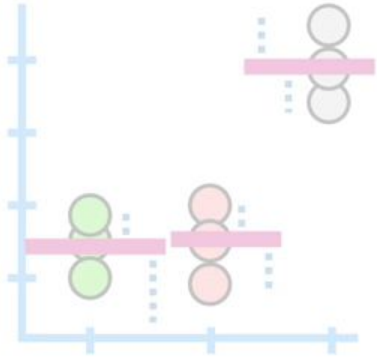
A decision about whether or not to ***reject*** or ***fail to reject*** the **Null Hypothesis**.

…we can end up making a ***different*** decision about the **Null Hypothesis**.

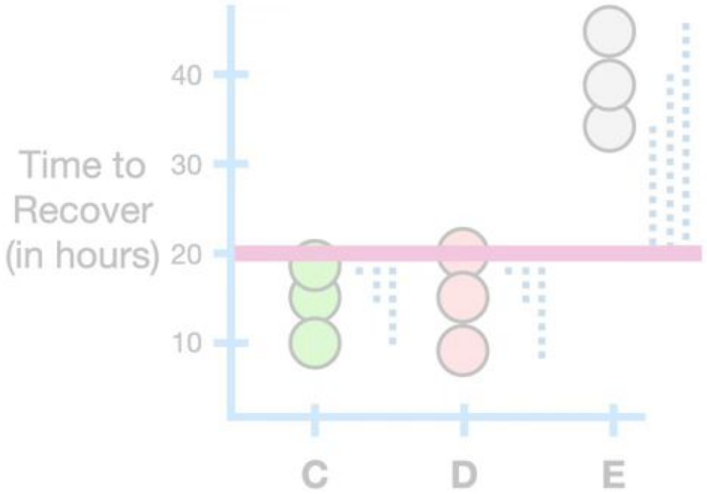
There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.



Drugs C, D are the same and **Drug E** is ***doing its own thing***.

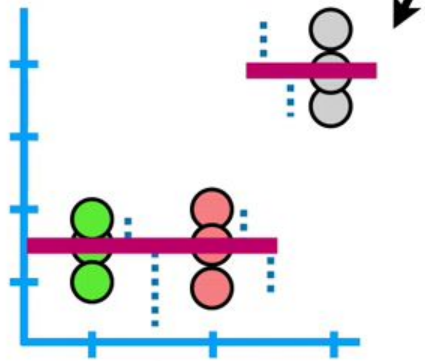


Drugs C, D and **E** are all different from each other.

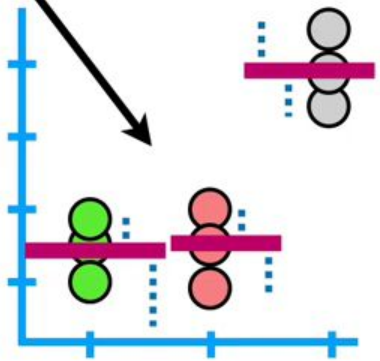


There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.

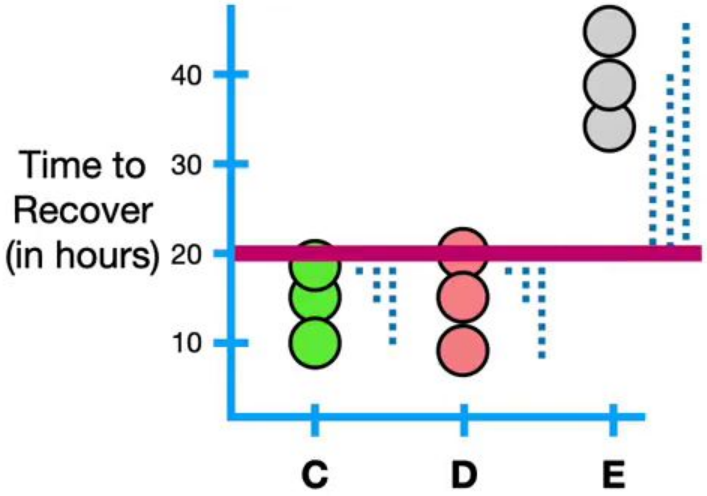
And that is why it is important to clearly state which alternative **Alternative Hypothesis** we want to use.



Drugs C, D are the same and **Drug E** is ***doing its own thing***.

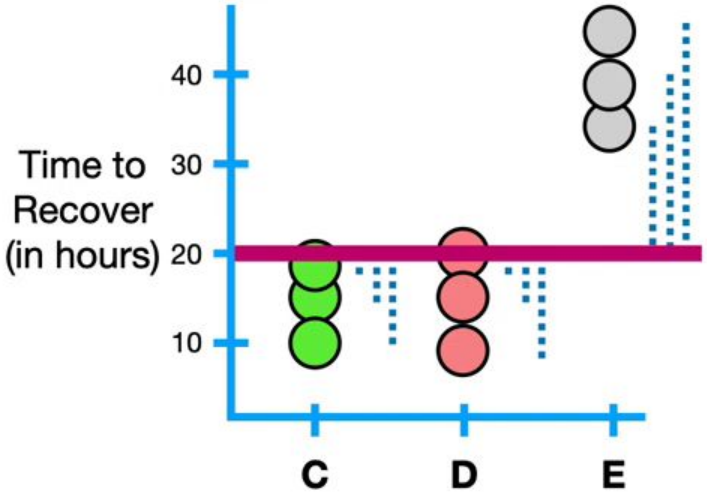


Drugs C, D and **E** are all different from each other.



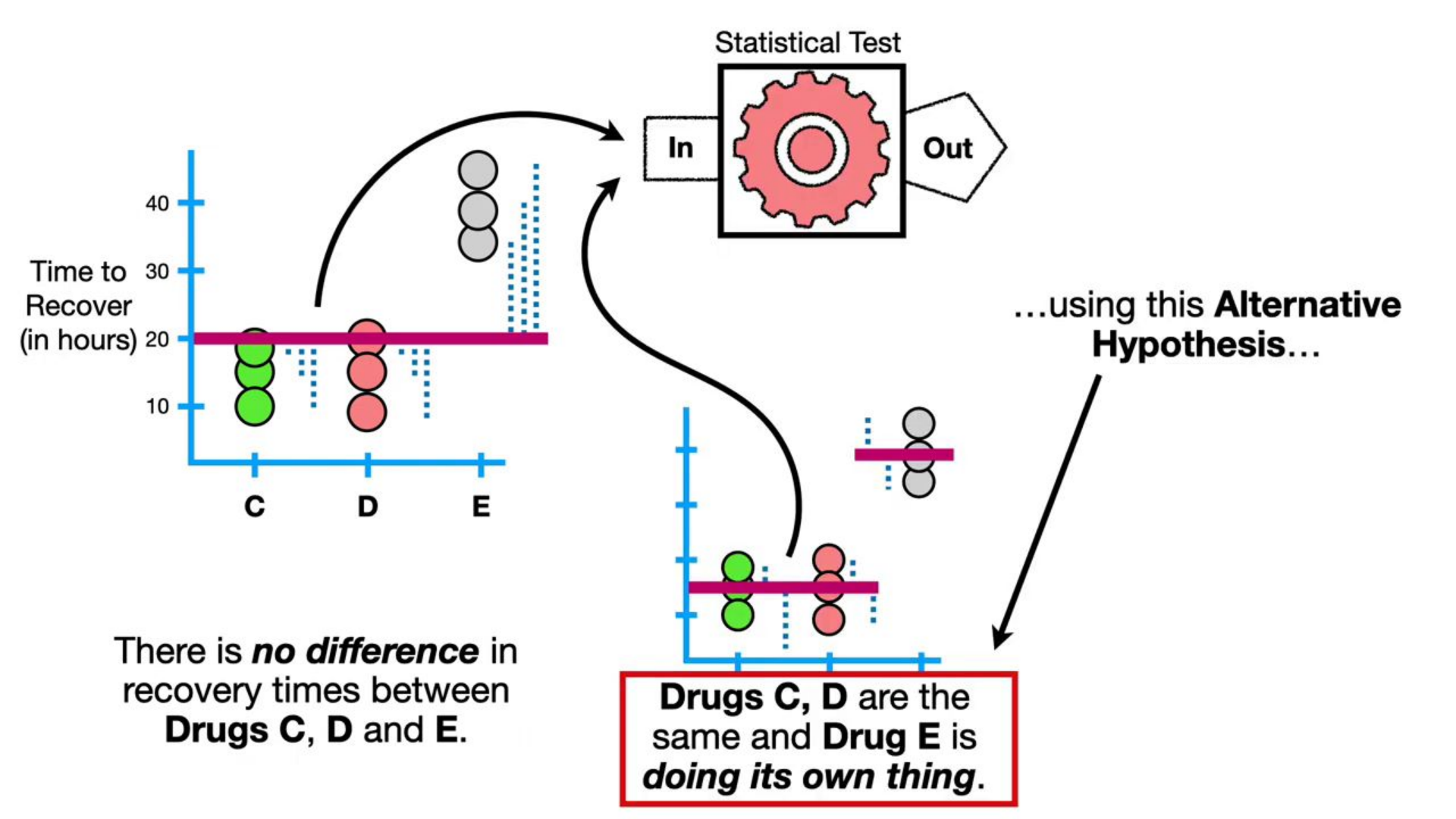
However, regardless of the **Alternative Hypothesis** we used in the test, we only ***reject*** or ***fail to reject*** the primary or **Null Hypothesis**.

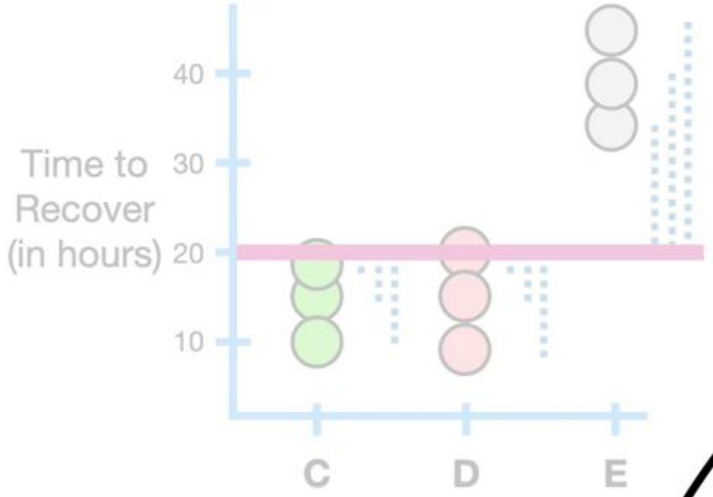
There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.



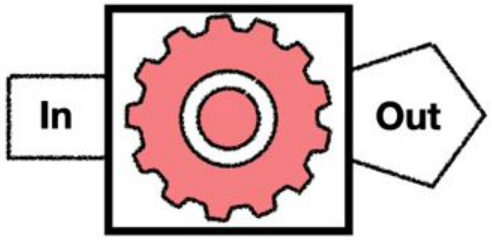
If we tested the **Null Hypothesis**…

There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.


Statistical Test
In
Out
40
30
20
10
Time to Recover (in hours)
C
D
E
There is *no difference* in recovery times between **Drugs C**, **D** and **E**.
…using this **Alternative Hypothesis**…
Drugs C, D are the same and **Drug E** is ***doing its own thing***.

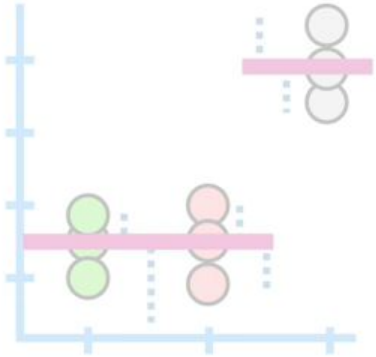


Statistical Test



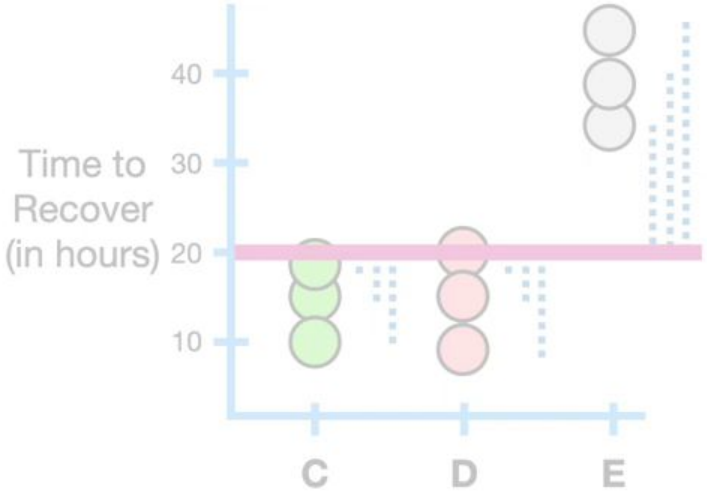
Reject the **Null Hypothesis.**

…and we ***rejected*** the **Null Hypothesis…**

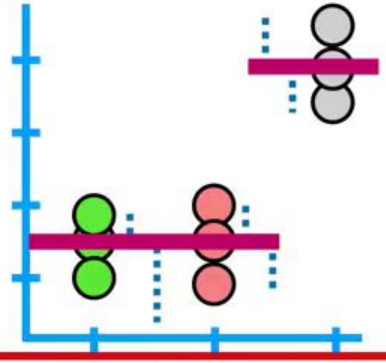


Drugs C, D are the same and **Drug E** is ***doing its own thing***.

~~There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.~~

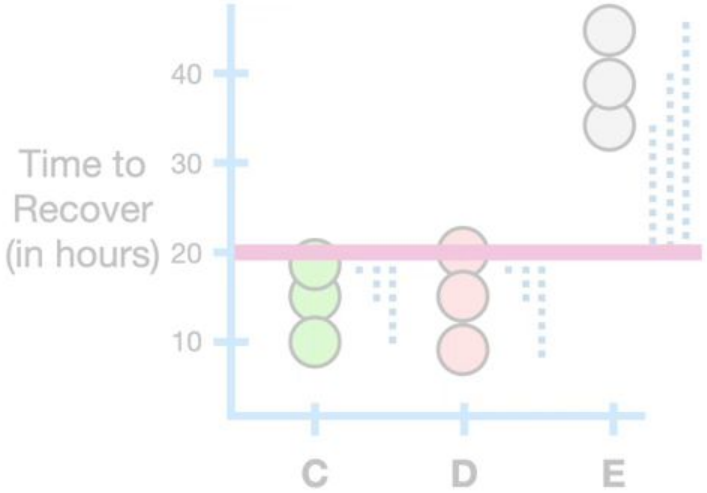


~~There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.~~

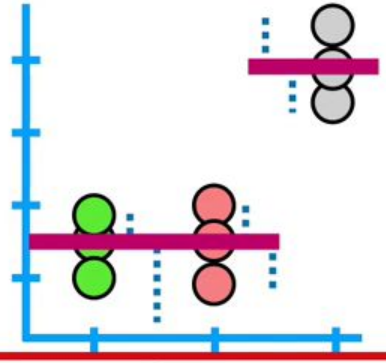


Drugs C, D are the same and **Drug E** is ***doing its own thing***.

…we might say that we rejected it in favor of this **Alternative Hypothesis**.

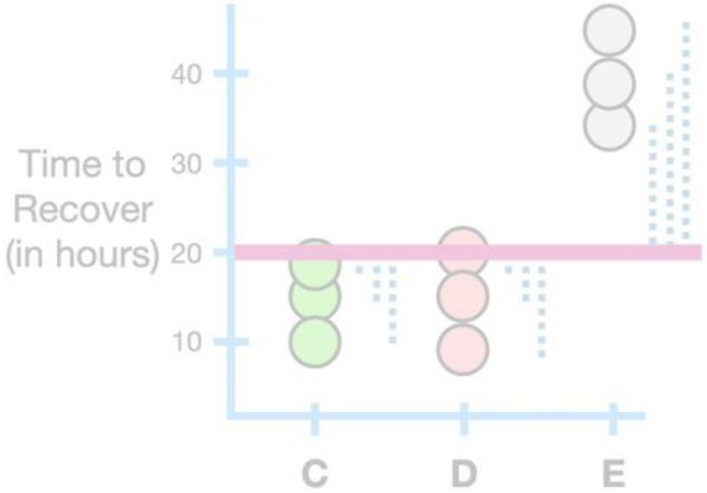


~~There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.~~

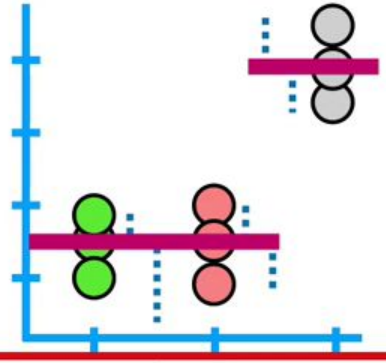


Drugs C, D are the same and **Drug E** is ***doing its own thing***.

However, we would still not say that we ***accept*** the **Alternative Hypothesis** because, just like we saw in the **StatQuest** on **Hypothesis Testing**, other **Alternatives** might be better.

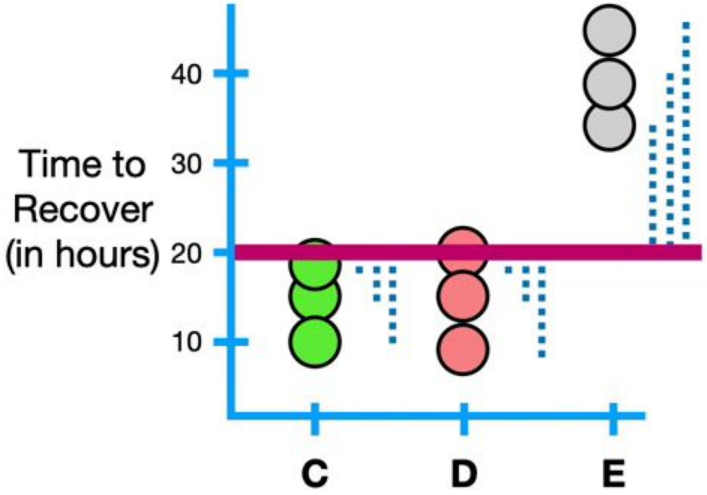


~~There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.~~



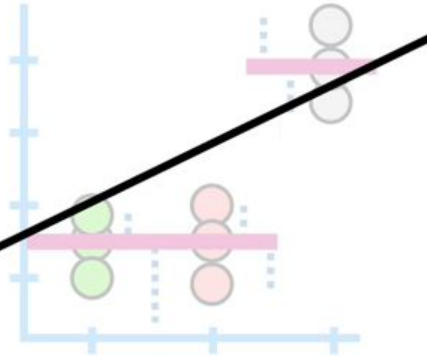
Drugs C, D are the same and **Drug E** is ***doing its own thing***.

In other words, there are too many possibilities to test to know if we have accepted the correct one.

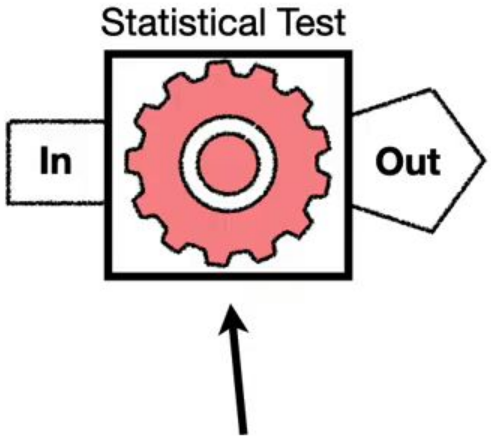


There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.

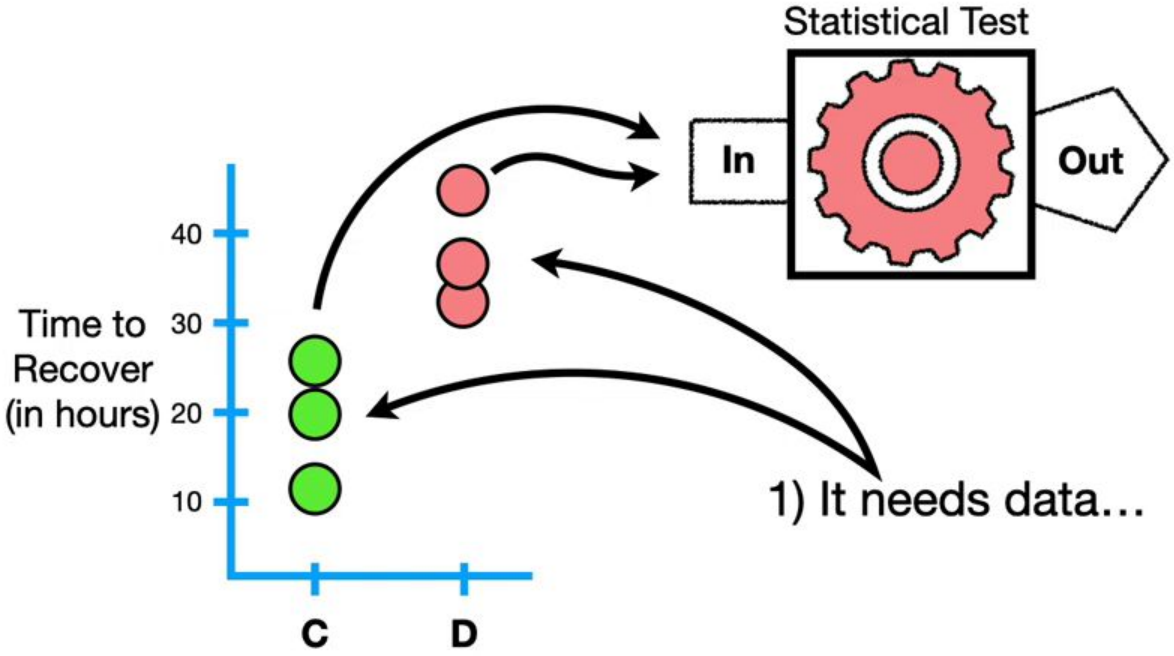
And is why we only ***reject*** or ***fail to reject*** the **null** or **primary hypothesis**.



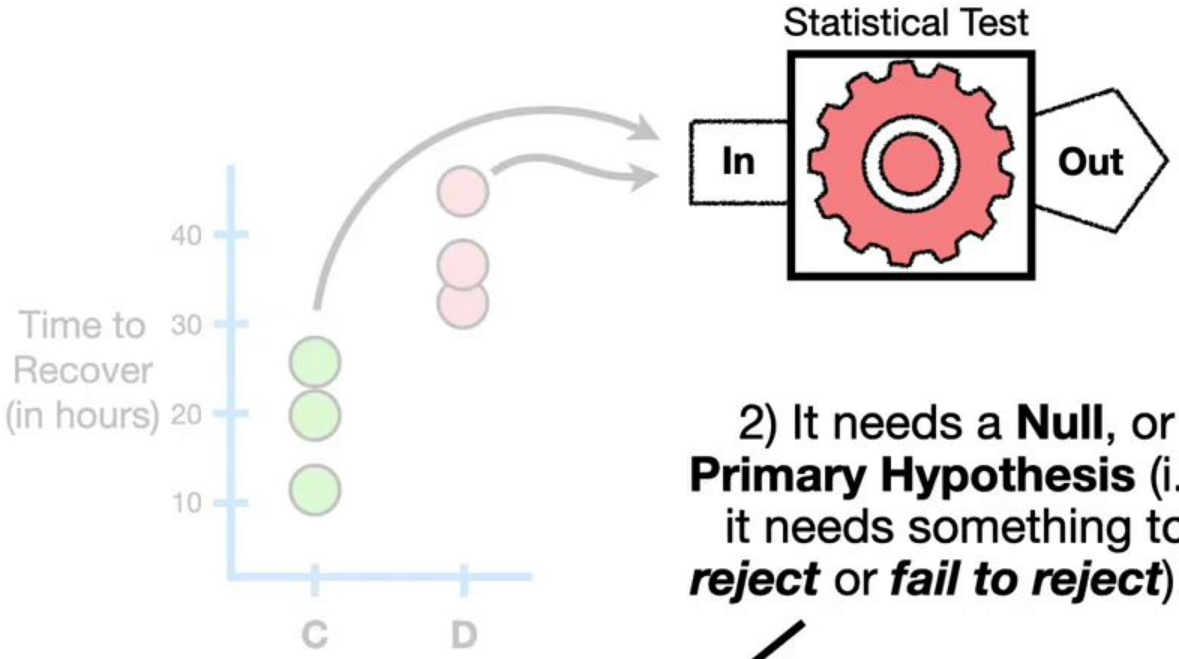
Drugs C, D are the same and **Drug E** is ***doing its own thing***.



A statistical test
needs **3** things.

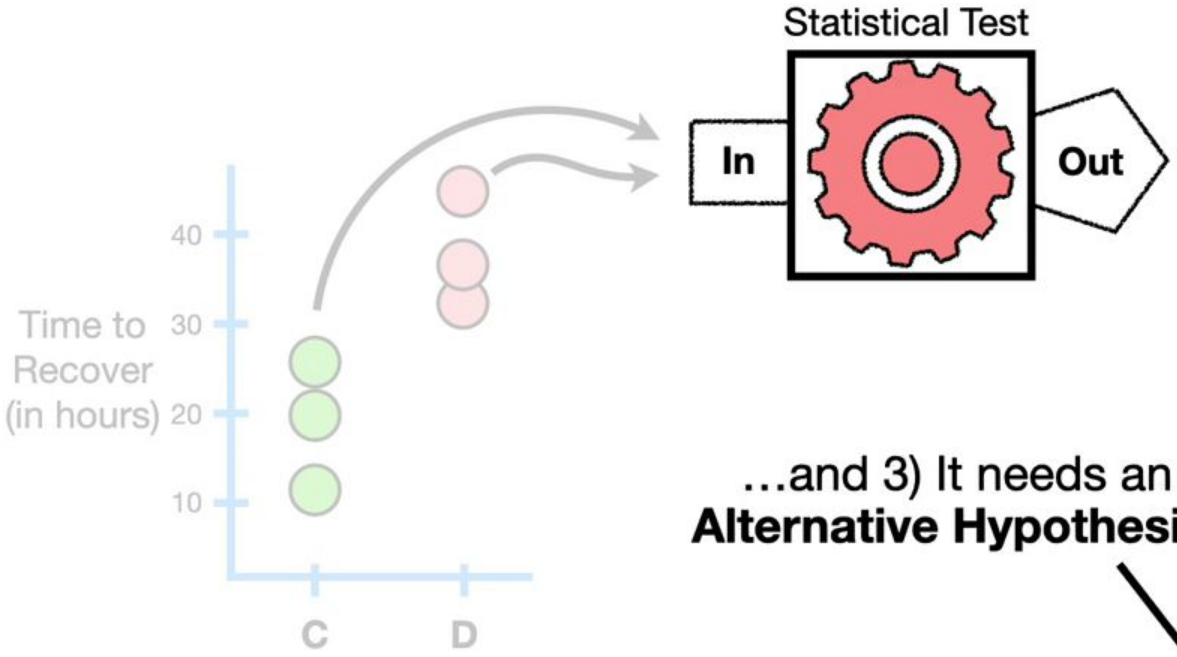


There is ***no difference*** in recovery times between **Drug C** and **Drug D**.



2) It needs a **Null**, or **Primary Hypothesis** (i.e. it needs something to ***reject*** or ***fail to reject***)…

There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

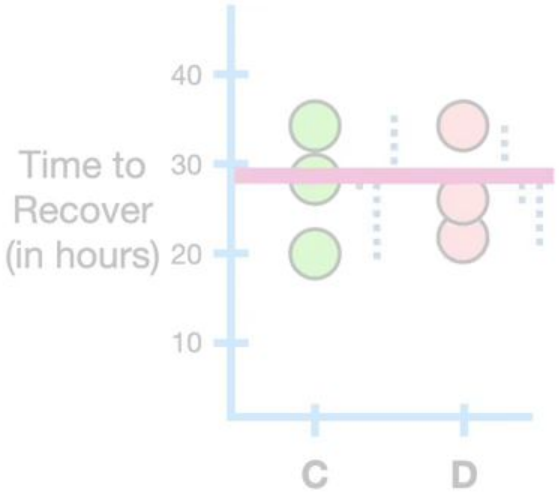


…and 3) It needs an **Alternative Hypothesis**.

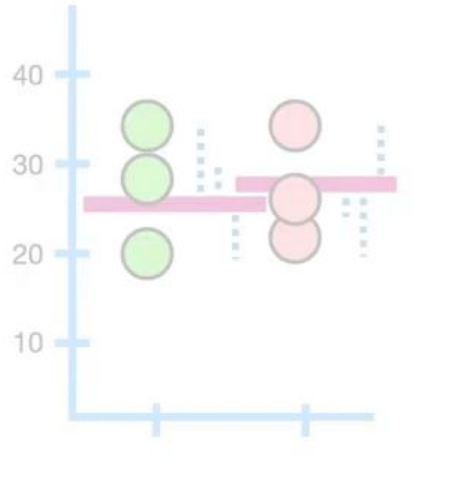
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.



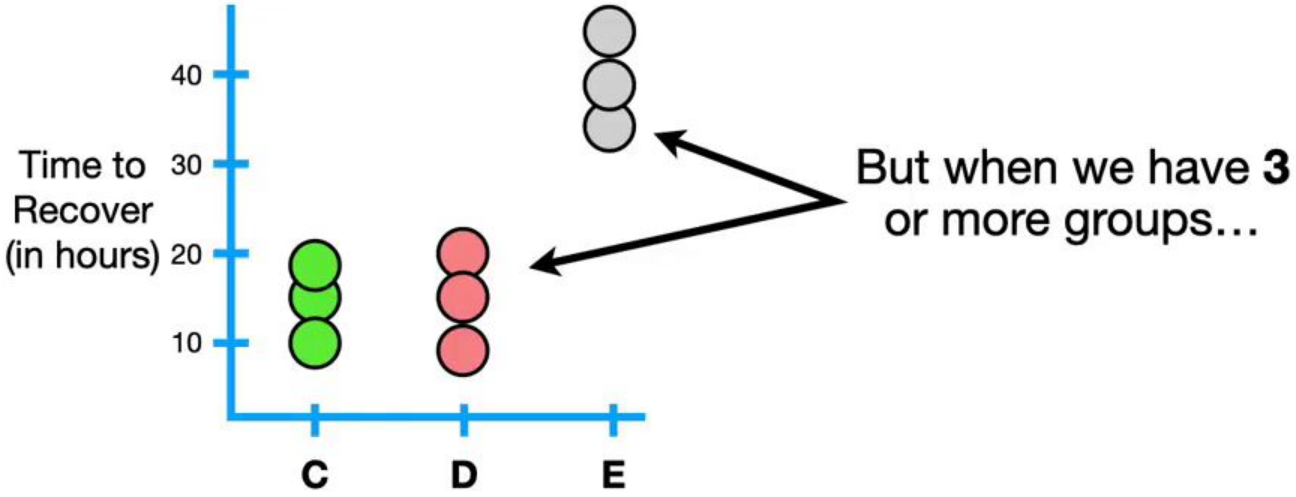
When we only have two groups of data, the **Alternative Hypothesis** is super obvious because it is just the opposite of the **Null Hypothesis**.



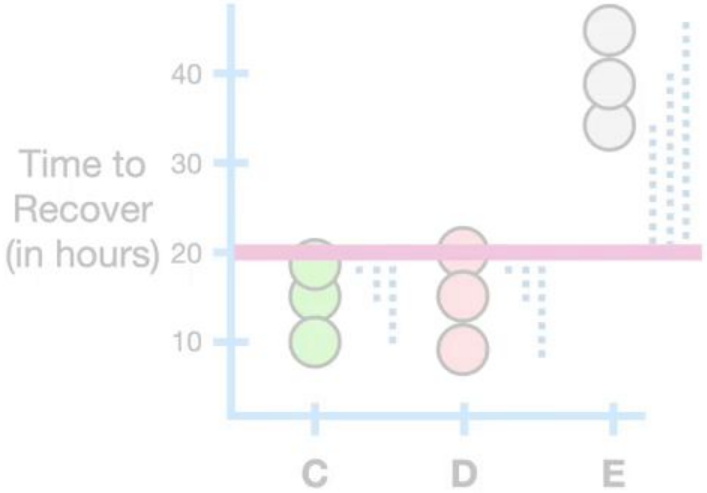
There is ***no difference*** in recovery times between **Drug C** and **Drug D**.

Alternative Hypothesis

There ***is a difference*** in recovery times between **Drug C** and **Drug D**.

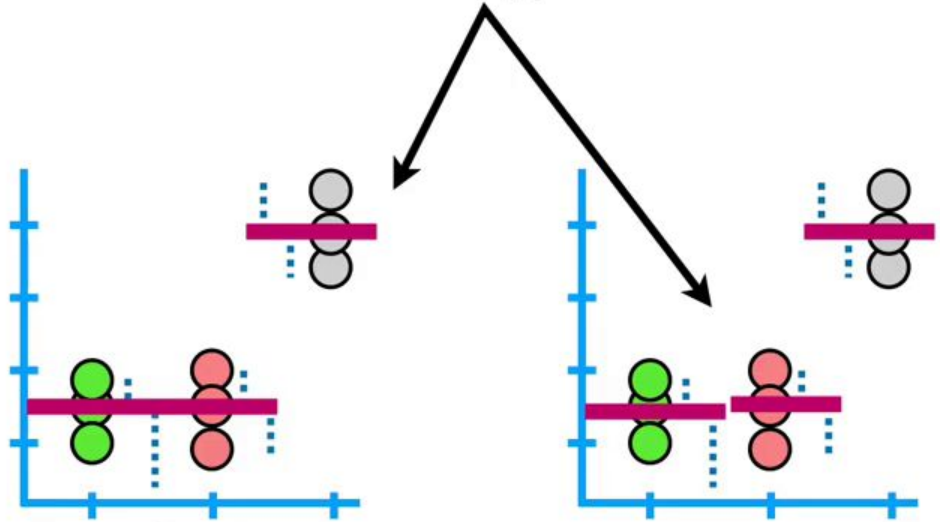


There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.



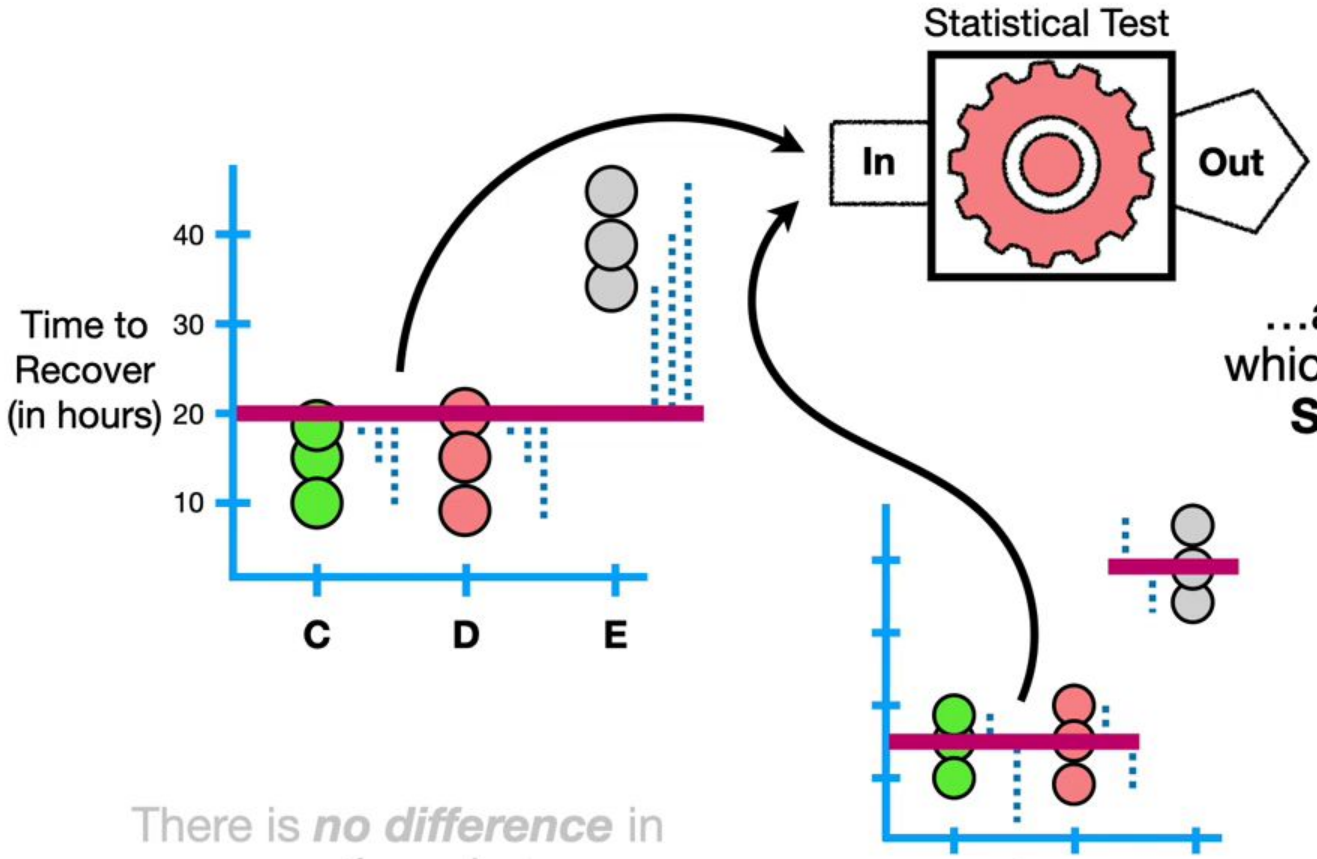
There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.

…we have options for the **Alternative Hypothesis.**



Drugs C, D are the same and **Drug E** is ***doing its own thing***.

Drugs C, D and **E** are all different from each other.

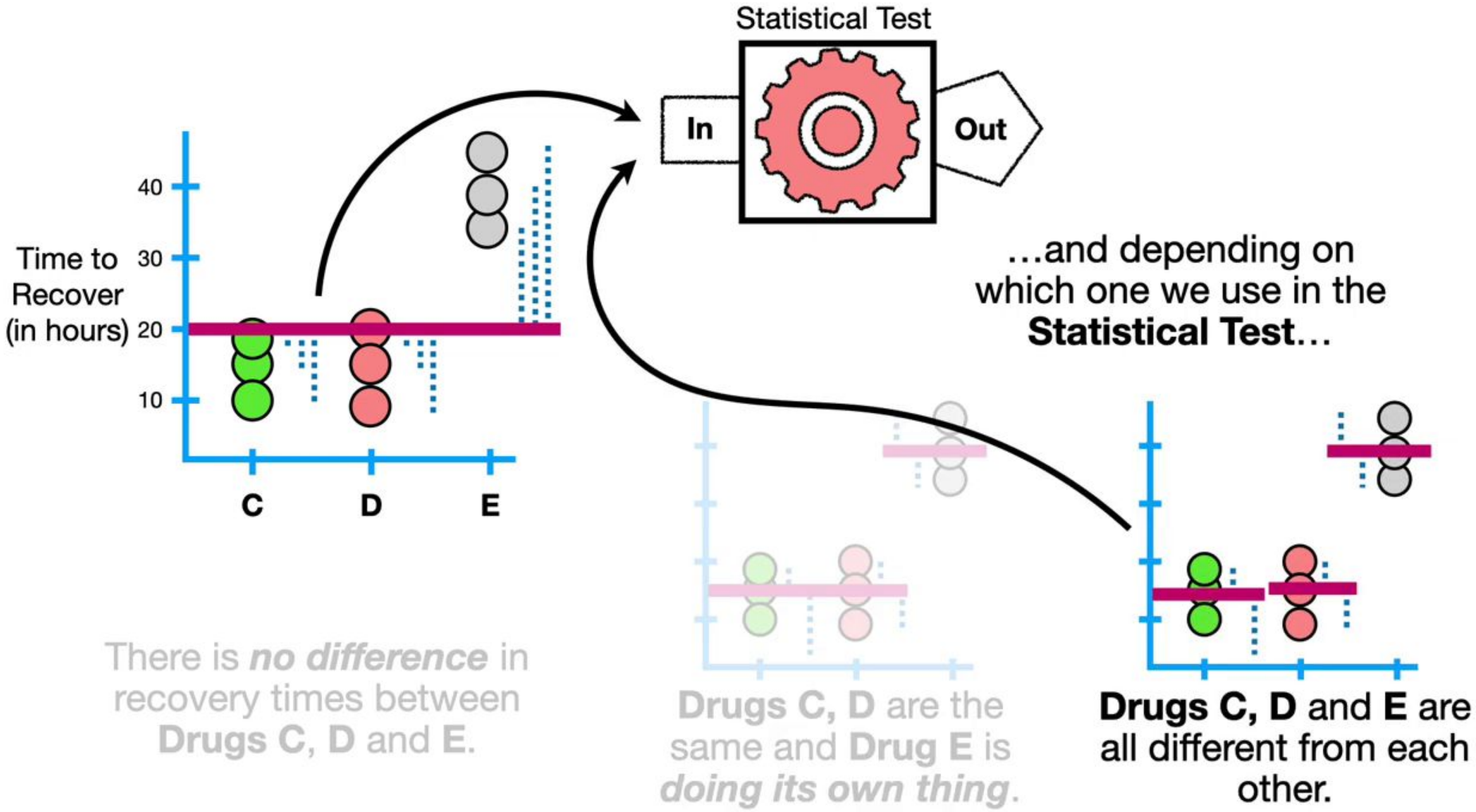


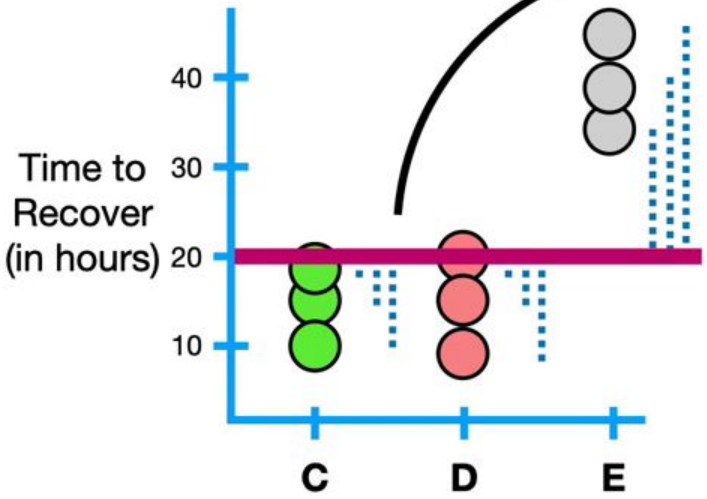
…and depending on which one we use in the **Statistical Test**…

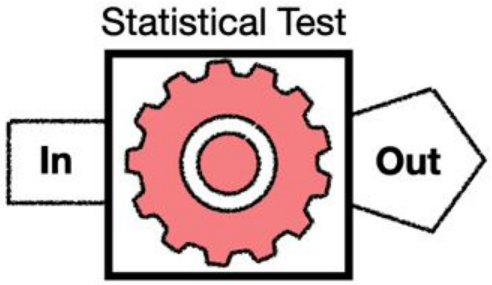
There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.

Drugs C, D are the same and **Drug E** is ***doing its own thing***.

Drugs C, D and **E** are all different from each other.


Statistical Test
In
Out
Time to Recover (in hours)
40
30
20
10
C
D
E
…and depending on which one we use in the **Statistical Test**…
There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.
Drugs C, D are the same and **Drug E** is ***doing its own thing***.
Drugs C, D and **E** are all different from each other.

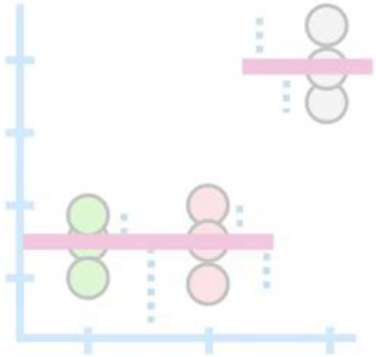




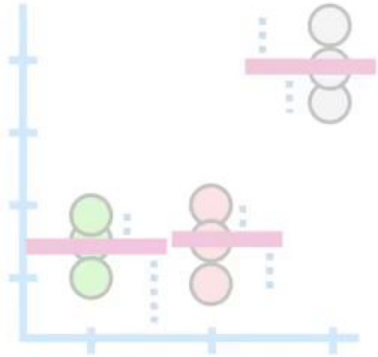
A decision about whether or not to ***reject*** or ***fail to reject*** the **Null Hypothesis**.

…we can end up making a ***different*** decision about the **Null**.

There is ***no difference*** in recovery times between **Drugs C**, **D** and **E**.



Drugs C, D are the same and **Drug E** is ***doing its own thing***.



Drugs C, D and **E** are all different from each other.