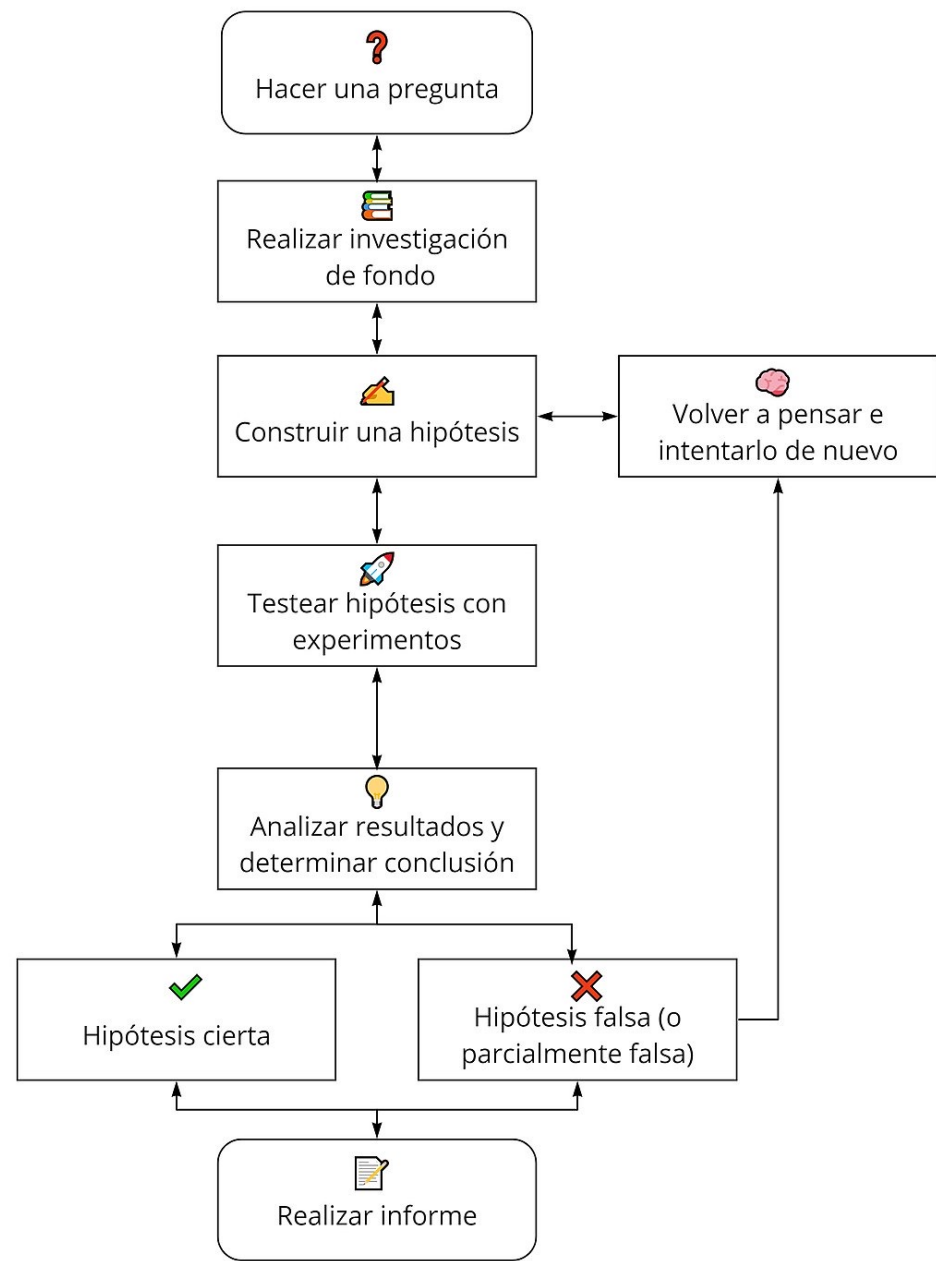


# **Diseño experimental y bioestadística**

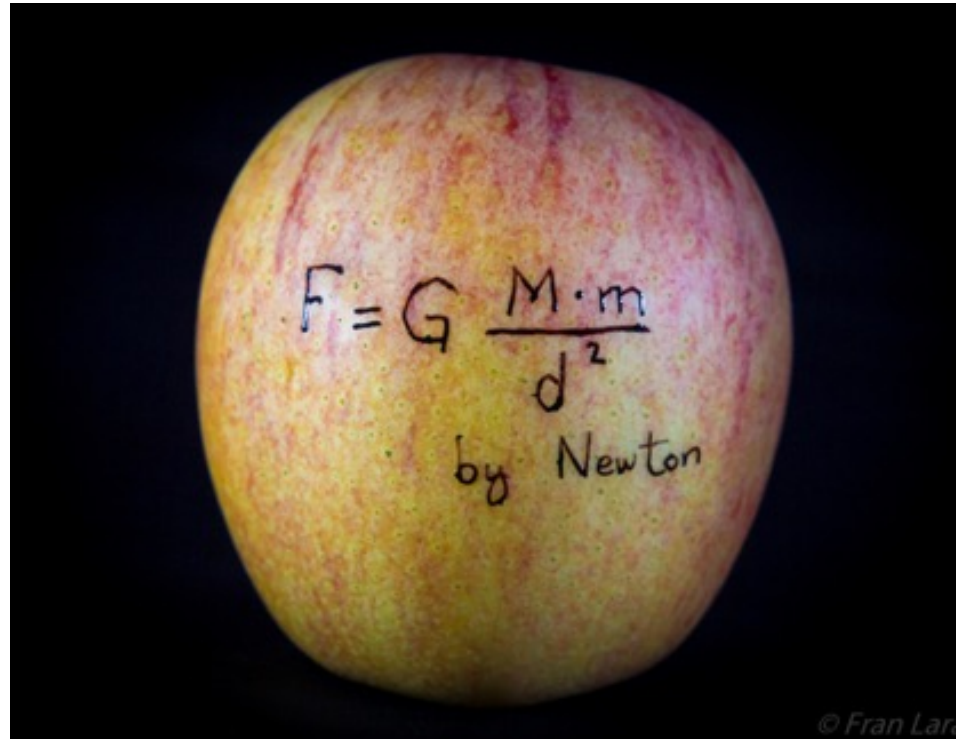
# Método científico

## Modelo simplificado de las etapas del método científico



# ¿Qué es una teoría?

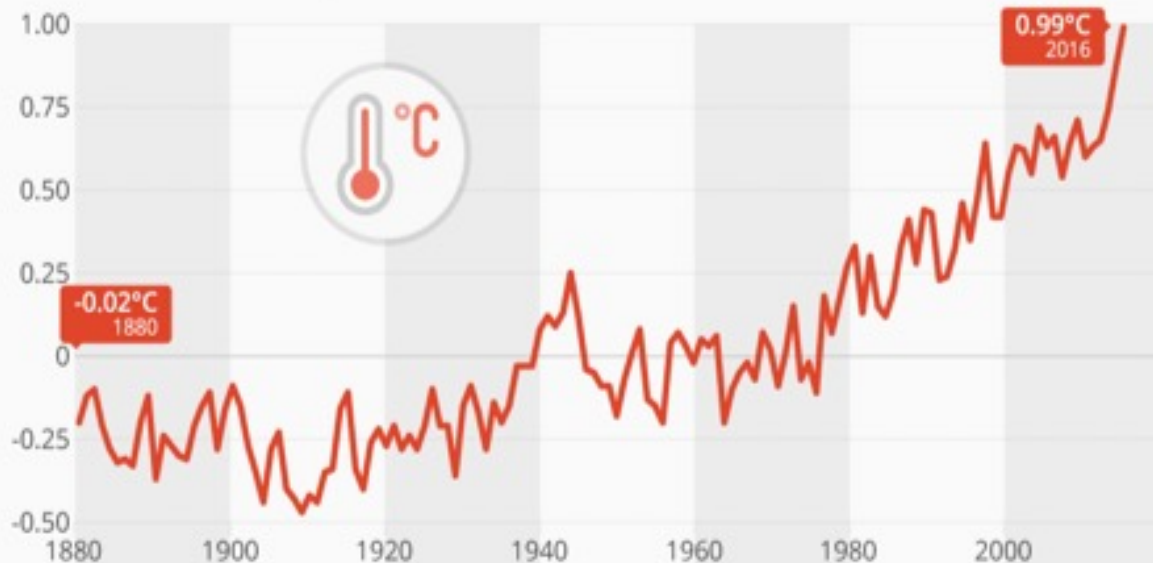
Es un conjunto organizado y coherente de ideas que explican un fenómeno biológico, las cuales surgen desde las observaciones, la experiencia y/o el razonamiento lógico (Futuyma & Kirkpatrick 2017, Wikipedia 2018).



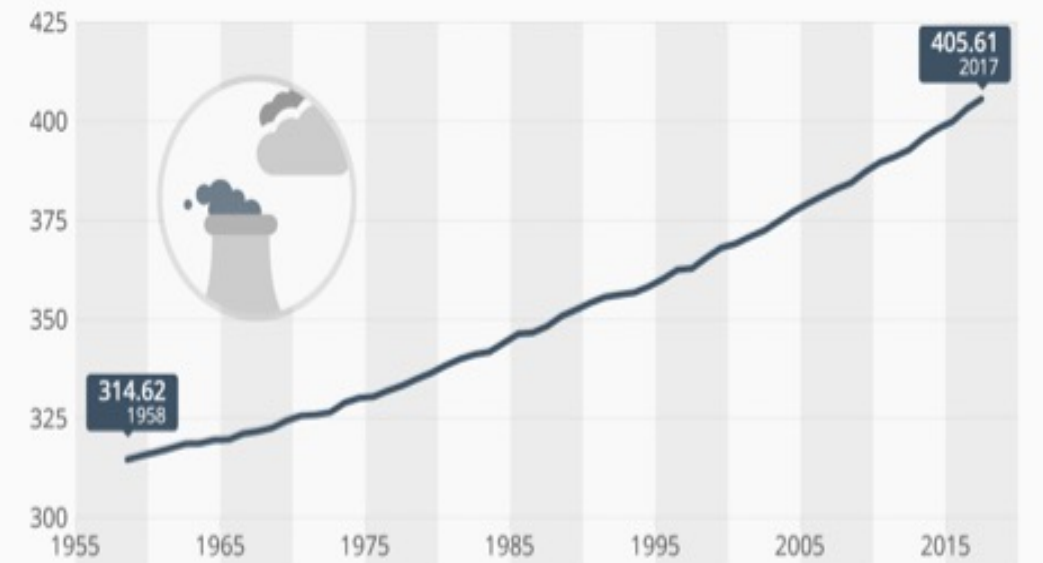
# ¿Qué es una hipótesis?

Es una explicación probable a un fenómeno natural (Futuyma & Kirkpatrick 2017).

Annual mean surface temperature of the earth from 1880 to 2016 (in °C)



Direct measurements of atmospheric carbon dioxide from 1958 to 2017 (in parts per million)\*



Hipótesis del calentamiento global de origen antropogénico

# Hipótesis

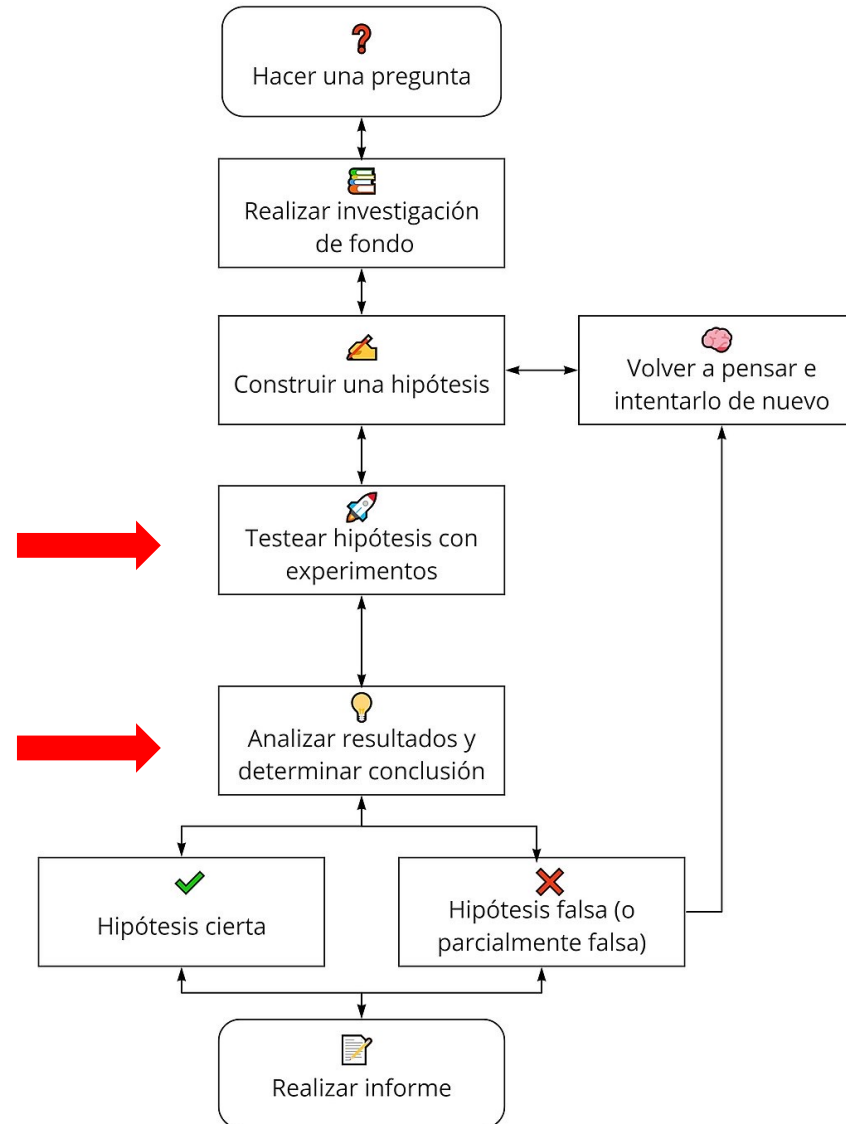
- Es una propuesta de cómo un sistema funciona.
- En base a las hipótesis se pueden hacer predicciones.
- Las predicciones se pueden poner a pruebas, ya sea realizando experimentos mensurativos o manipulativos.
- Si los resultados son consistentes con la la predicción, luego la hipótesis es aceptada.



- La hipótesis es una formulación forma de una idea de cómo funciona un sistema.
- Sin embargo, aprobar o rechazar hipótesis, no necesariamente implica un avance científico.
- Responder preguntas si implican avance, llenando un vacío en el conocimiento.

# Método científico

Modelo simplificado de las etapas del método científico



*Diseño experimental*

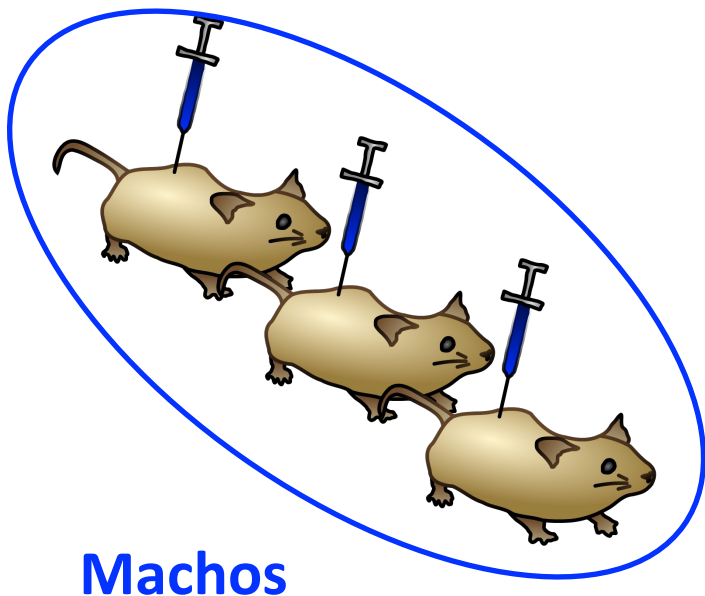


*Análisis estadístico*

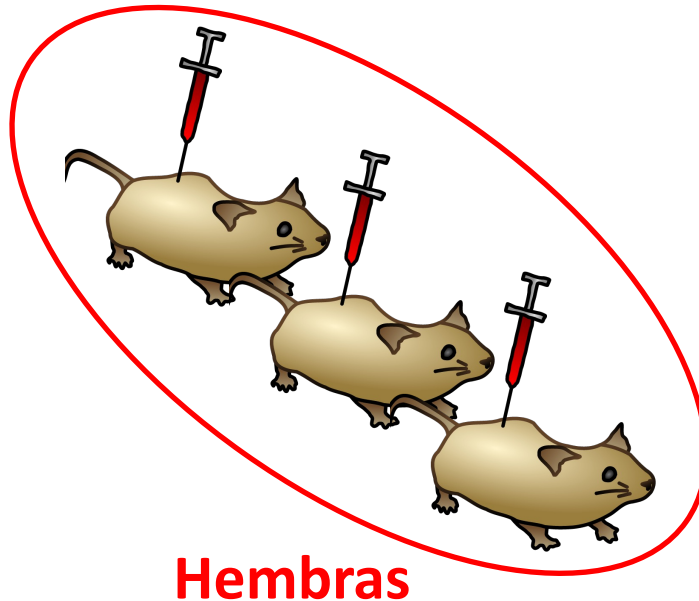
# Diseño experimental

- Un diseño experimental apropiado debe minimizar el efecto de factores o variables de confusión.
- **Los factores o variables de confusión** restringen la interpretación inequívoca de un experimento, ya que las diferencias entre los tratamientos experimentales no pueden ser separadas de otros factores que podrían causar las diferencias observadas.

## Efecto de una hormona sobre la agresividad en ratones



Machos



Hembras

El efecto de la hormona y del sexo no pueden ser separados entre sí.

Ambos factores se confunden entre sí.

# Diseño experimental

- Las variables o factores confundentes son tan evidentes y pueden bastante comunes, incluso en trabajo publicados en revistas científicas.
- Las principales fuentes de confusión en el diseño experimental se deben a:
  - Replicación inapropiada.
  - Falta de controles apropiados.
  - Falta de asignación apropiada de las unidades experimentales a los distintos experimentos.

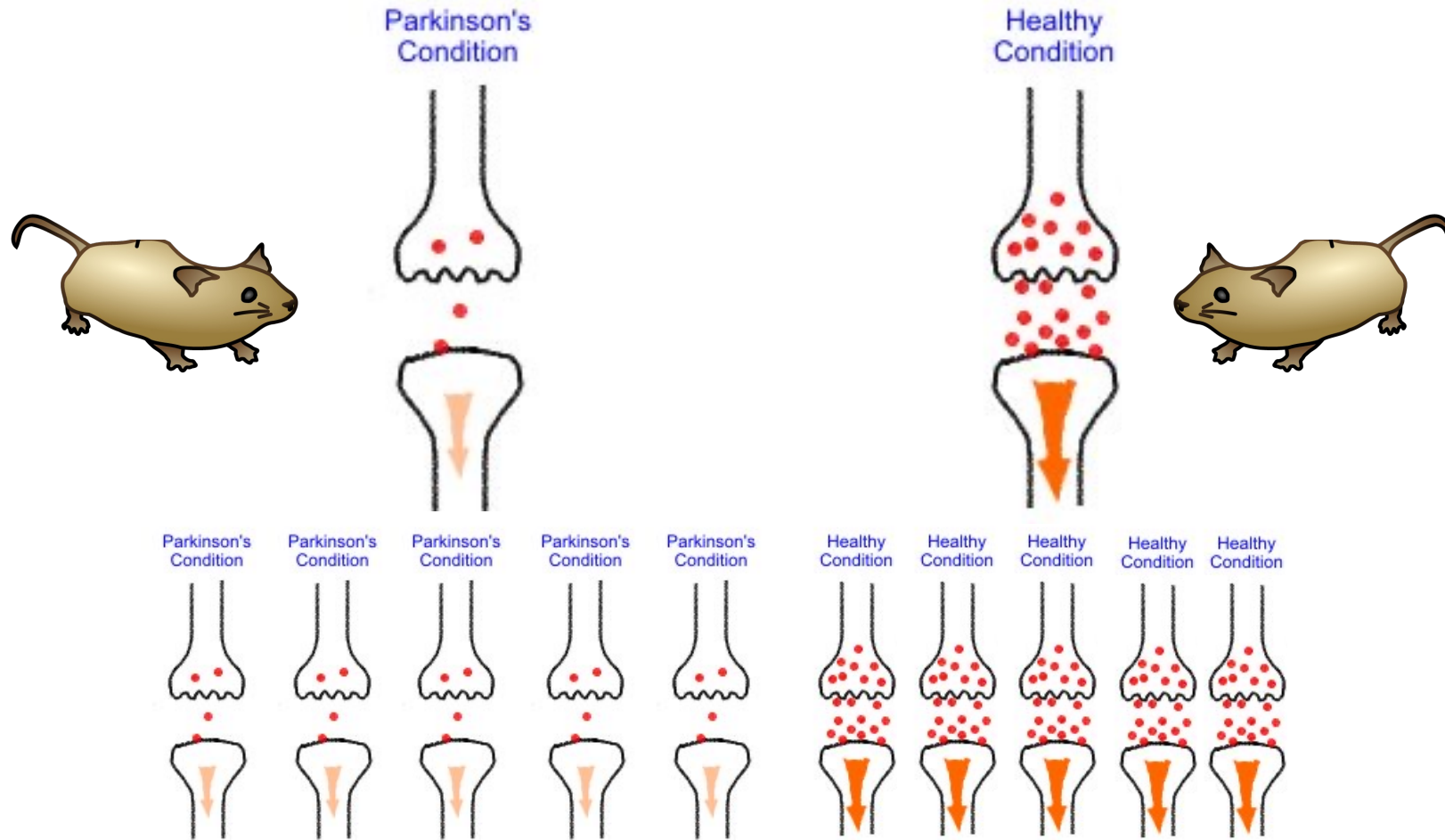
## Replicación

- La replicación se refiere a tener observaciones replicadas a un escala espacial y/o temporal que concuerde con la aplicación de los tratamientos experimental.
- La replicación es necesaria porque los sistemas biológicos son inherentemente variables.



# Diseño experimental

## Niveles de dopamina entre ratones controles y ratones con mutaciones de riesgo para Parkinson



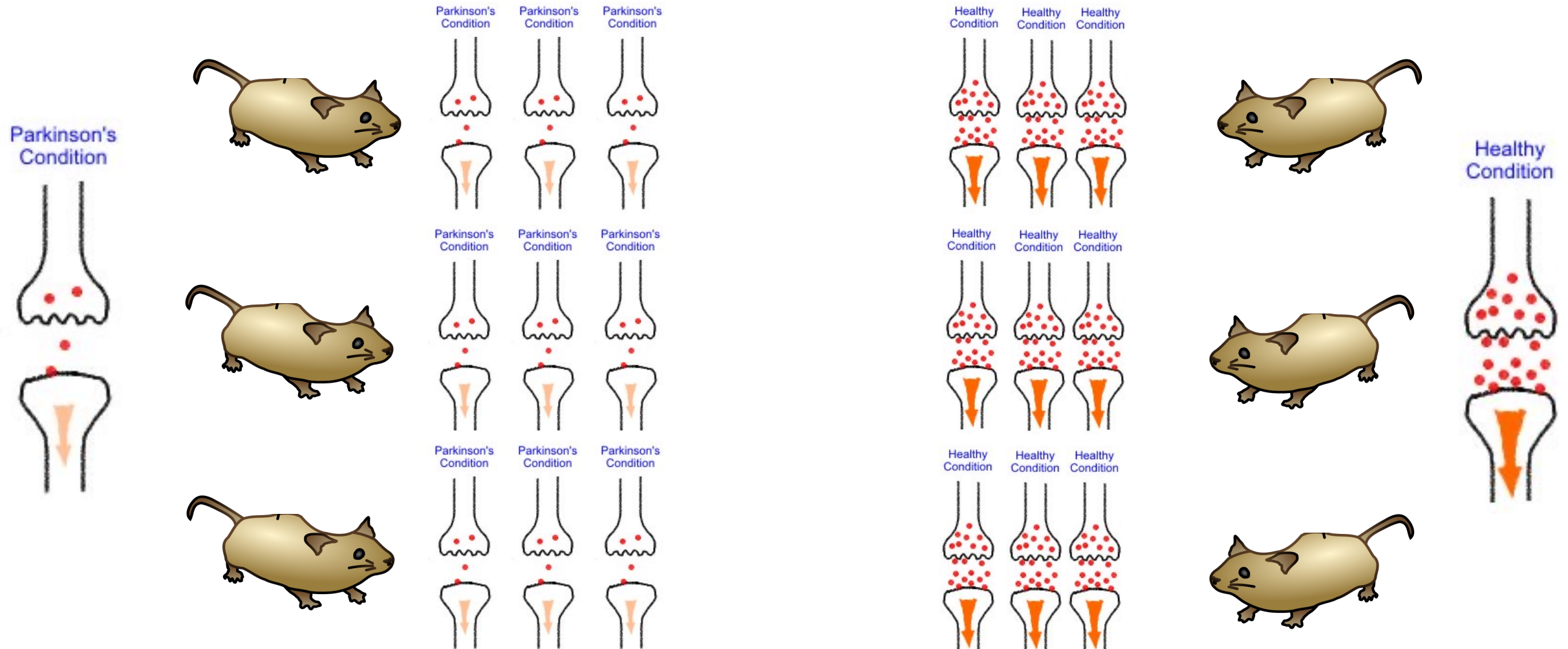
Las unidades experimentales no son las neuronas, sino los ratones.

# Diseño experimental

- Existe solo un ratón como replica experimental para cada una de las condiciones: ratón control versus ratón con mutaciones de riesgo para Parkinson.
- El efecto de estar o no afectado se confunde completamente con las diferencias inherentes entre las dos ratones.
- Si existen diferencias, solo podemos concluir que hay diferencias entre las dos ratones.
- Las diferentes neuronas analizadas al pertenecer a un mismo ratón (misma unidad experimental) son consideradas submuestras. Esto es lo que se conoce como **pseudoreplicación**.
- En este caso, el diseño se encuentra replicado, pero en la escala incorrecta.
- Una opción es aumentar el número de unidades experimentales (ratones) para aumentar el nivel de replicación.
- Sin embargo, hay que tener cuidado con la pseudoreplicación (submuestras asociadas a unidades experimentales).

# Diseño experimental

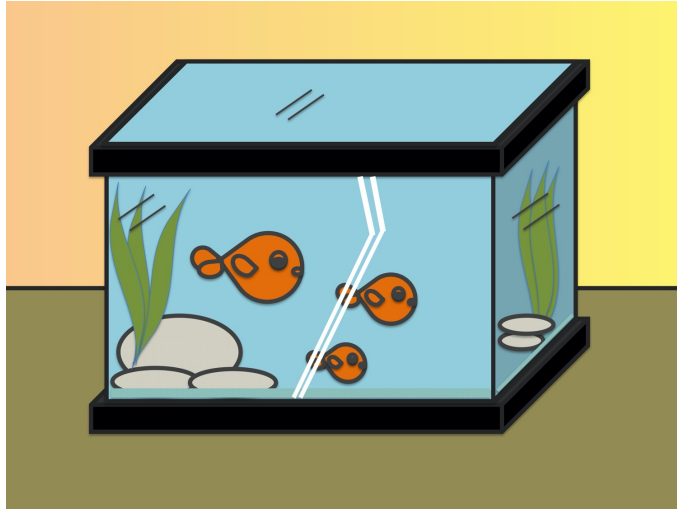
## Niveles de dopamina entre ratones controles y ratones con mutaciones de riesgo para Parkinson



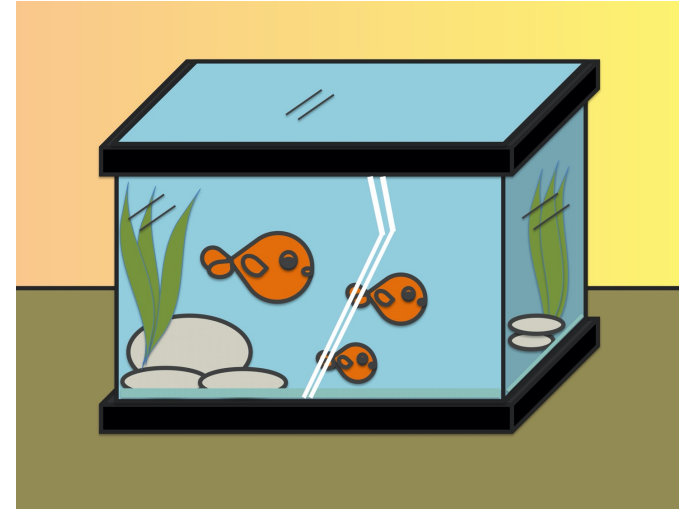
# Diseño experimental

**Efecto de la exposición a metales pesados sobre la expresión génica en un vertebrado marino**

Control



Cobre



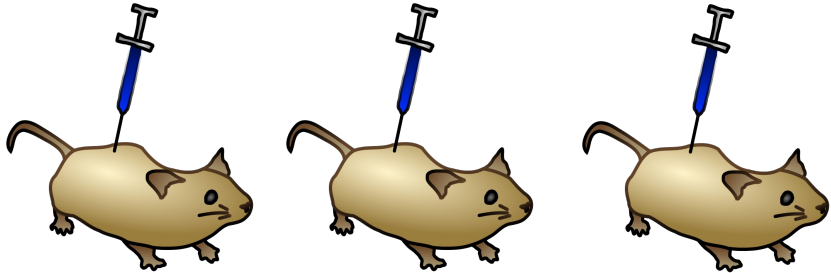
Este diseño experimental tiene control, pero bajo nivel de replicación.

# Diseño experimental

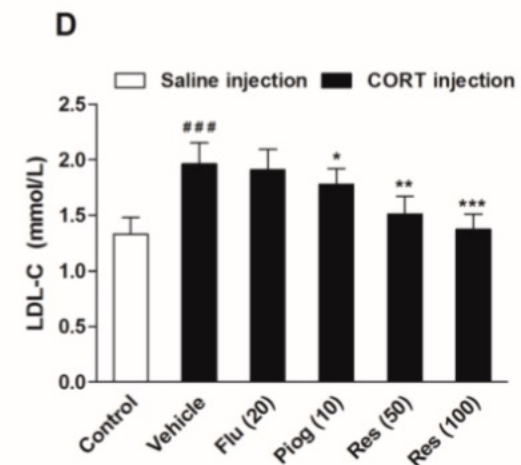
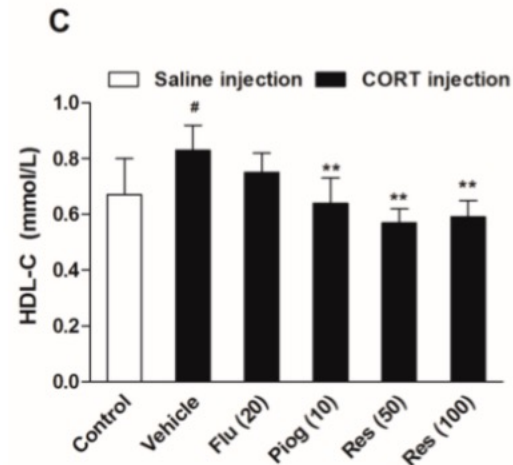
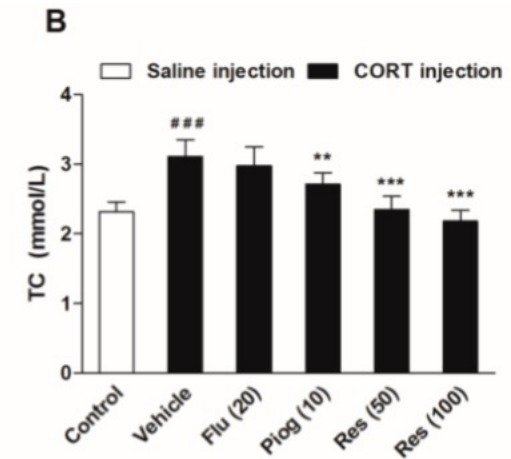
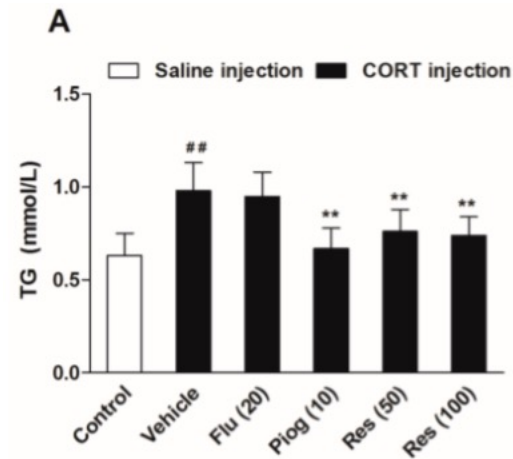
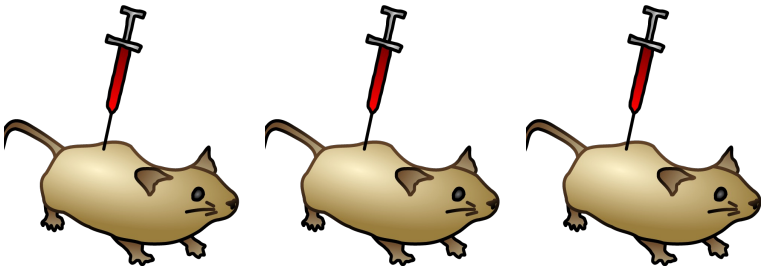
## Controles

- La función de los controles es esencial para evaluar qué sucede cuando no hay manipulación experimental.

### Control



### Tratamiento



# Diseño experimental

## Randomización (aleatorización)

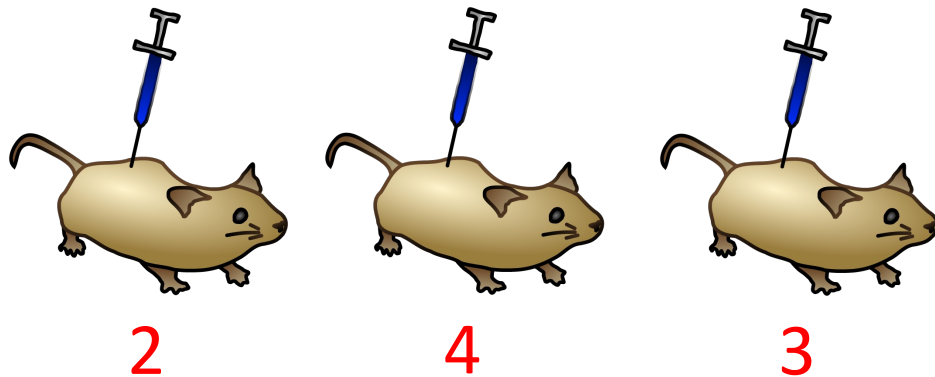
- Las muestras representan unidades experimental de una población.
- A partir de las muestras podemos estimar parámetros poblacionales como la media, la varianza, etc.
- Las unidades experimentales deben representar una muestra aleatoria de todas las posibles unidades experimentales.
- Si, las muestras analizadas realmente representan a la población, las variables respuestas deberían ser independientes de las unidades experimentales que son analizadas.
- Por ejemplo, los ratones utilizados en el experimento anterior deberían representar una muestra aleatoria de todos los posibles ratones que podrían haber sido incorporados en el experimento.

# Diseño experimental

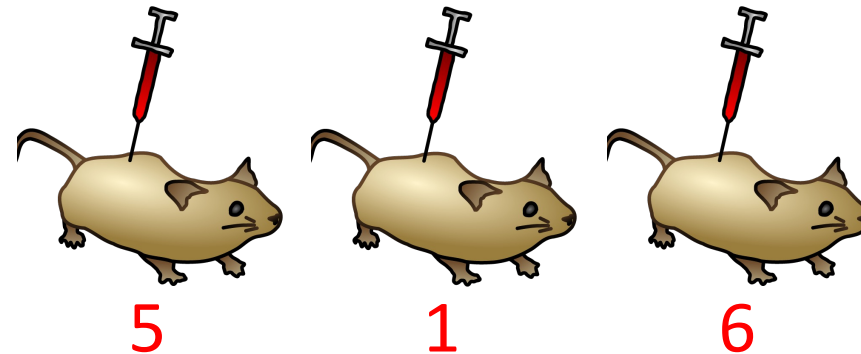
## Randomización (aleatorización)



## Control

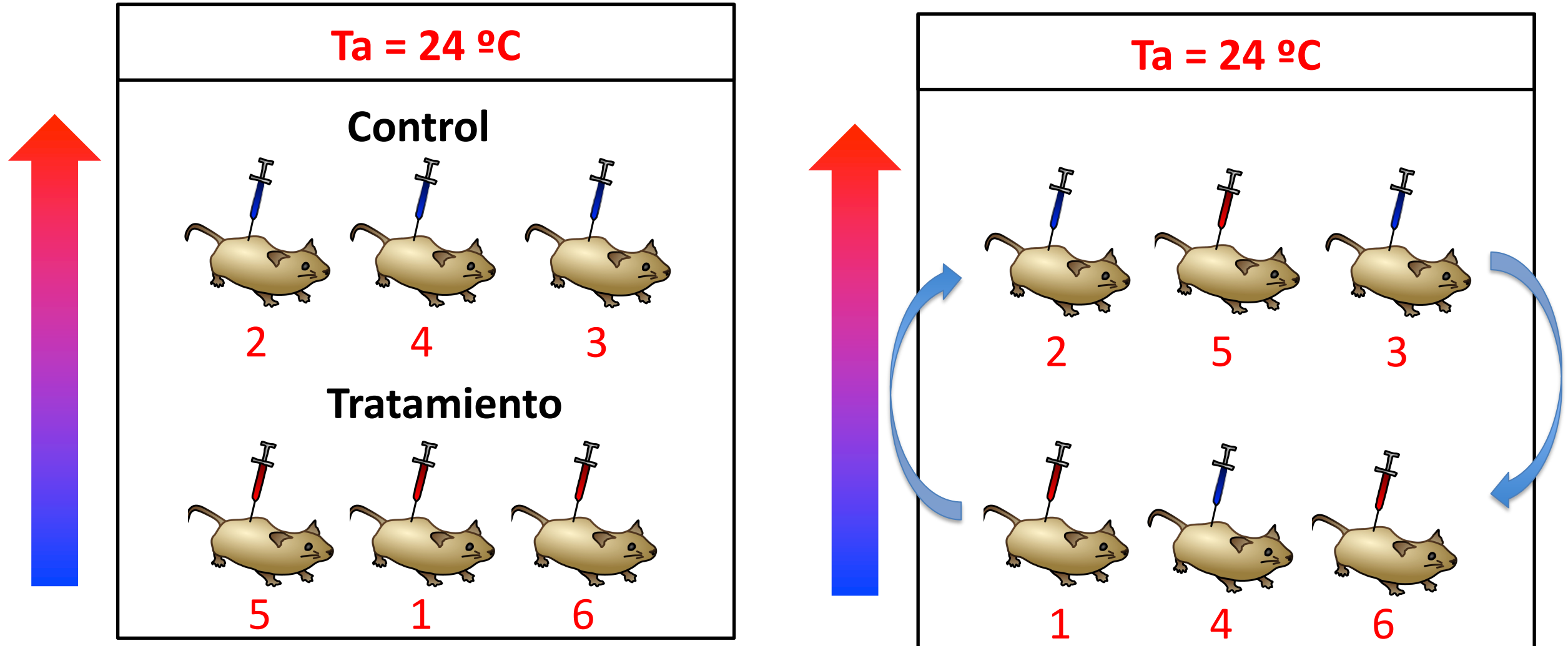


## Tratamiento



# Diseño experimental

## Randomización (aleatorización)



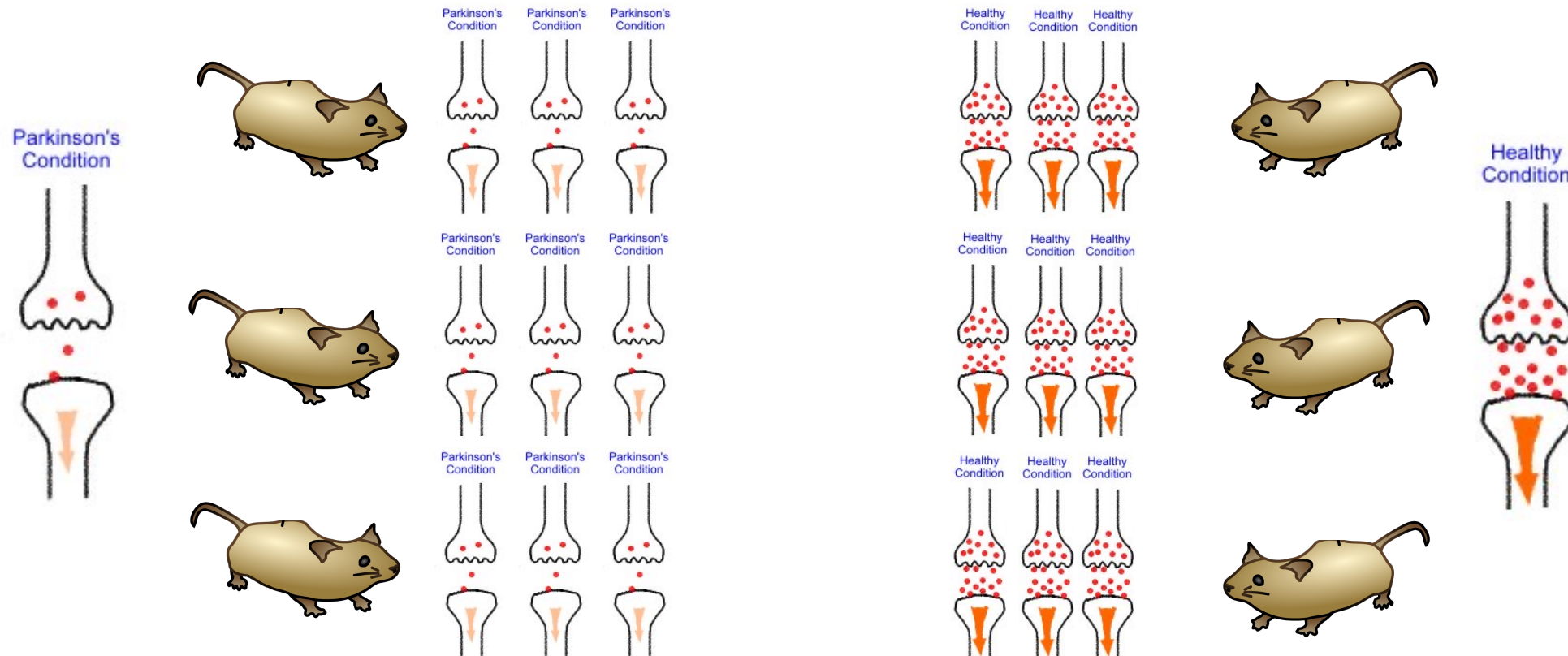


# Módulo 3: Diseño experimental

## Independencia

- La falta de independencia entre las muestras puede dificultar la interpretación de los resultados e invalidar la aplicación de varios análisis estadísticos.

## Niveles de dopamina entre ratones controles y ratones con mutaciones de riesgo para Parkinson



# Diseño experimental

## Poder estadístico

- Es una medida de nuestra confianza de haber detectado efectos importantes en nuestro experimento si estos realmente existieron.
- Esta asociado a la capacidad de detectar un efecto en nuestras muestras si estas realmente existen en la población.
- En el caso de no encontrar diferencias entre nuestros tratamientos, cabe preguntarse si nuestro diseño experimental tenía la capacidad de detectar las diferencias que posiblemente existían entre los tratamientos.
- Realizar análisis de poder también nos permitiría justificar la existencia de resultados no significativos.

$$Poder = \frac{TE * \alpha * \sqrt{n}}{\sigma}$$

*El poder aumenta a mayor tamaño del efecto (TE) y número de replicas (n), y disminuye con la variabilidad ( $\sigma$ ) entre muestras.*

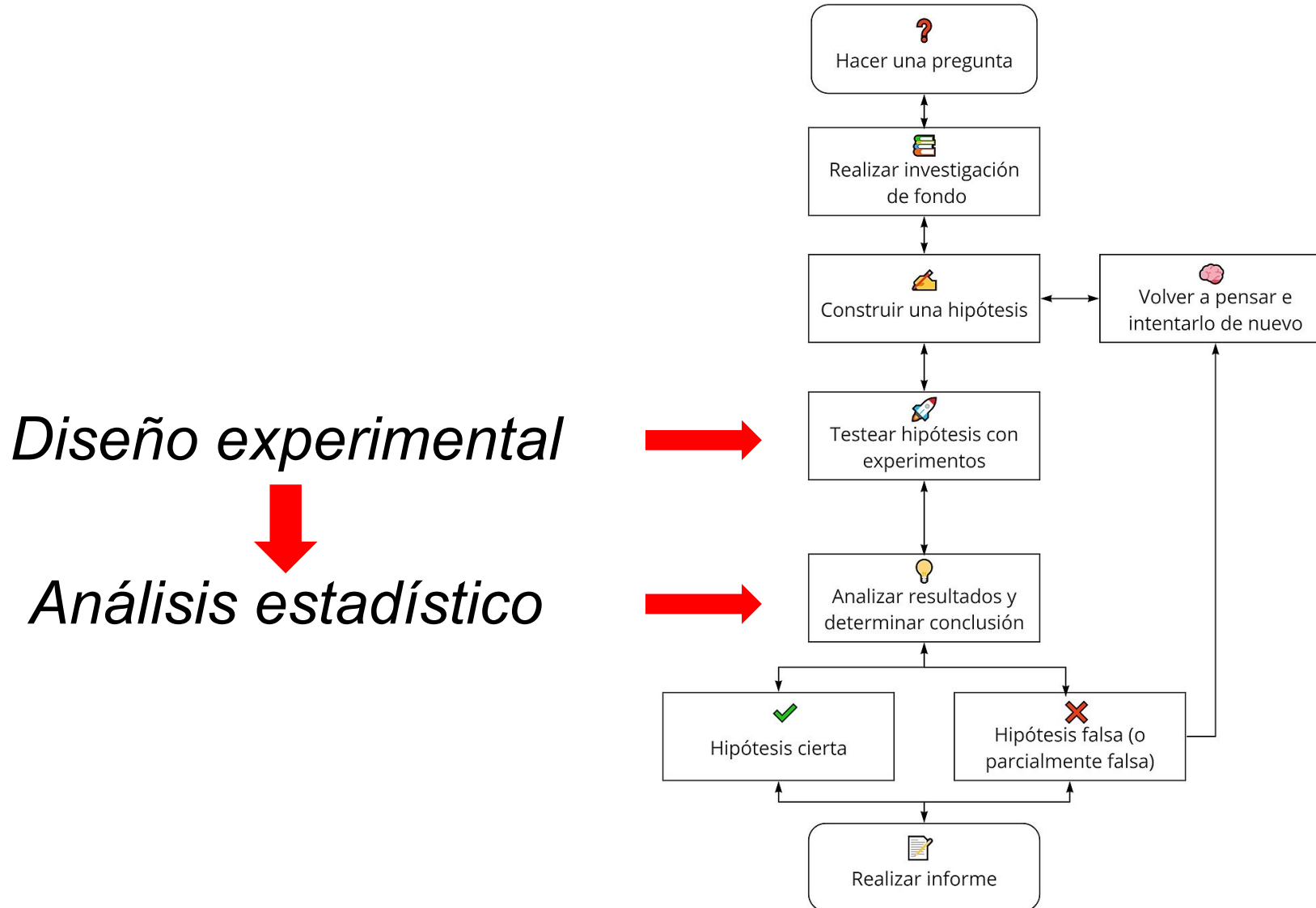
# Diseño experimental

## Resumen

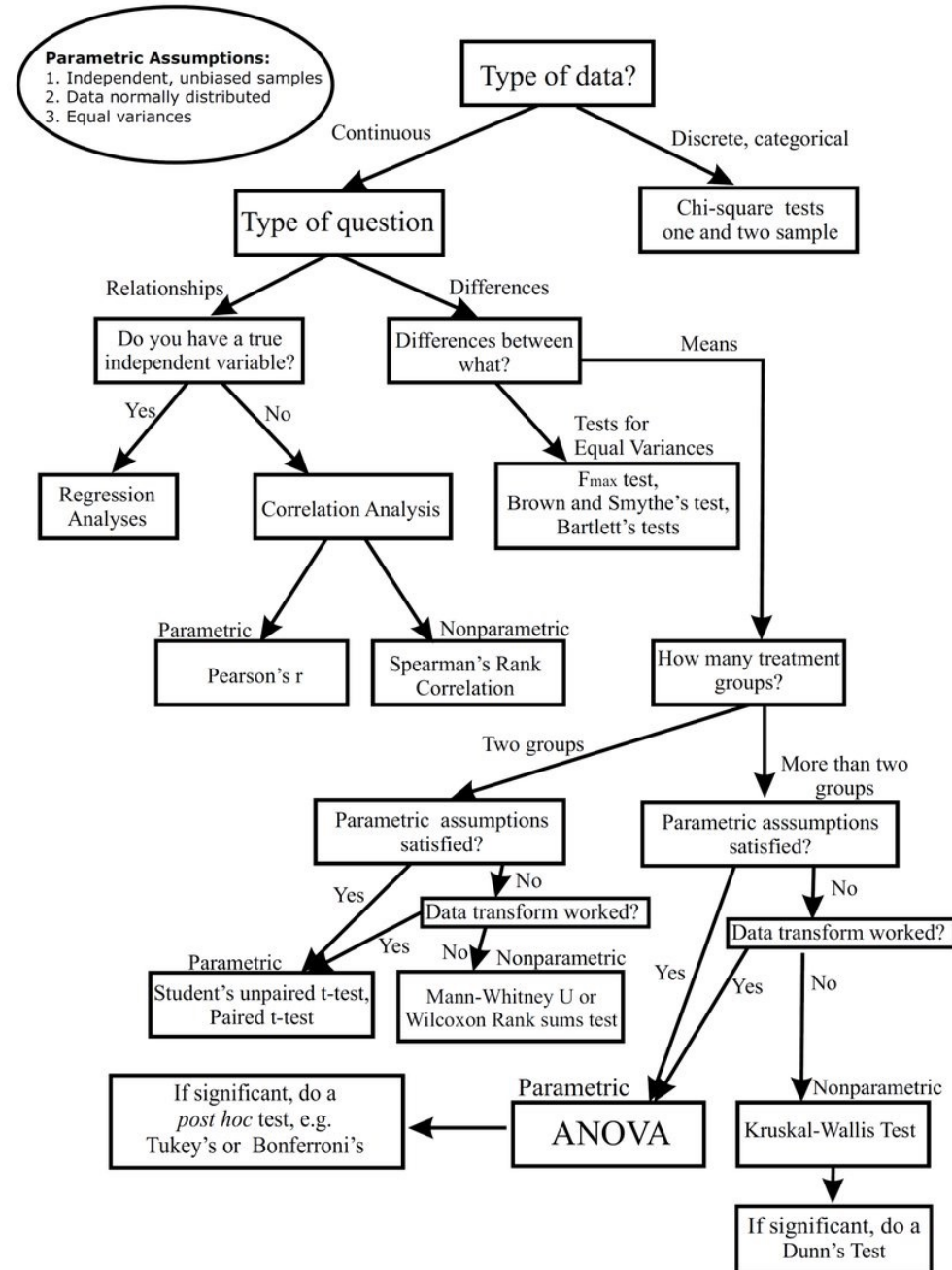
- Establecer claramente las respuestas esperadas y no esperadas. Es decir, plantear claramente las hipótesis y/o preguntas de investigación.
- Al momento de diseñar el experimento, evaluar críticamente la validez de los controles y las unidad experimentales.
- Identificar el modelo estadístico a aplicar para analizar los datos.
- Identificar los supuestos de los análisis estadísticos
- Idealmente realizar un experimento piloto para estimar la variación de la variable respuesta
- Realizar análisis de poder para evaluar si nuestro diseño experimental es adecuado y realizable (tamaño muestral alcanzable).
- Discutir y pensar críticamente y de forma grupal el diseño experimental.
- Una vez decidido nuestro diseño experimental, no hay vuelta atrás. Nuestra conclusiones estarán restringidas a la capacidad de nuestro diseño para poner a prueba nuestras hipótesis.

# Método científico

Modelo simplificado de las etapas del método científico

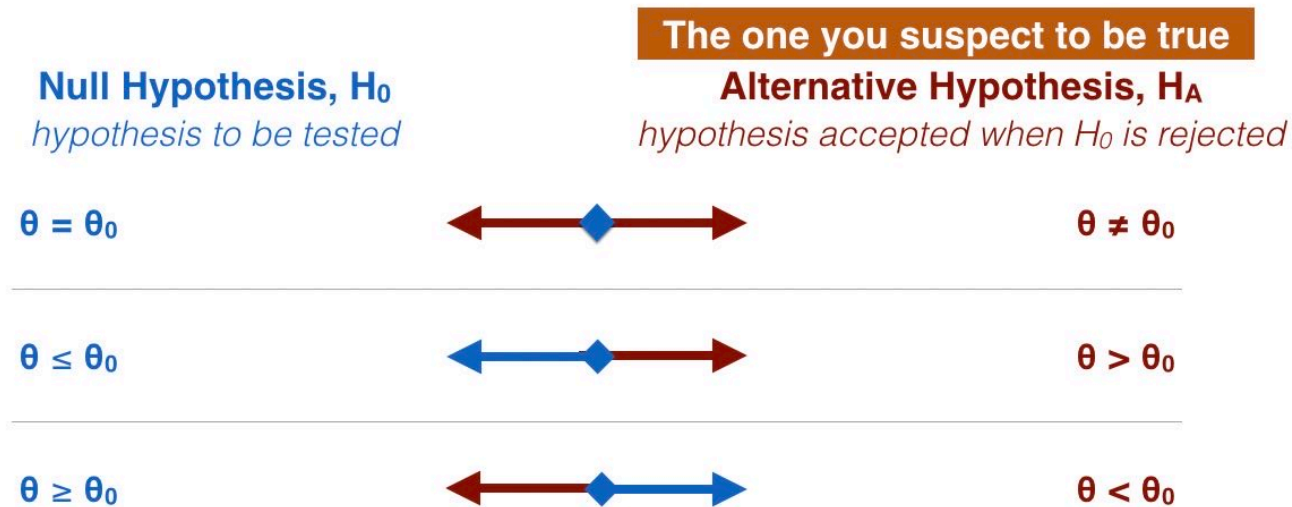


# Flow Chart for Selecting Commonly Used Statistical Tests



# Prueba de hipótesis

- Es una propuesta de cómo un sistema funciona.
- En base a las hipótesis se pueden hacer predicciones.
- Las predicciones se pueden poner a pruebas, ya sea realizando experimentos mensurativos o manipulativos.
- Si los resultados son consistentes con la la predicción, luego la hipótesis es aceptada.



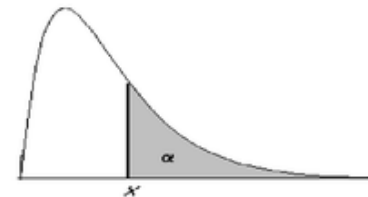
**Nivel de significancia ( $\alpha$ )** es la probabilidad de rechazar la hipótesis nula cuando es verdadera (error tipo I).

Ronald Fisher (1925) propuso que un  $\alpha = 0.05$  era un adecuado nivel de certeza.

Si  $\chi^2$  calculado es mayor al  $\chi^2$  de tabla según  $\alpha$  y GL,  $H_0$  es rechazada.

Tabla de la distribución chi-cuadrado.

La tabla contiene los valores  $x$  tales que  $p[\chi_n^2 \geq x] = \alpha$  en función de los grados de libertad ( $n$ ).



| $n$ | 0,99   | 0,98   | 0,975  | 0,95   | 0,90   | 0,80   | 0,50   | 0,20    | 0,10    | 0,05    | 0,025   | 0,02    | 0,01    | 0,001   |
|-----|--------|--------|--------|--------|--------|--------|--------|---------|---------|---------|---------|---------|---------|---------|
| 1   | 0,0002 | 0,0006 | 0,0010 | 0,0039 | 0,0158 | 0,0642 | 0,4549 | 1,6424  | 2,7055  | 3,8415  | 5,0239  | 5,4119  | 6,6349  | 10,8274 |
| 2   | 0,0201 | 0,0404 | 0,0506 | 0,1026 | 0,2107 | 0,4463 | 1,3863 | 3,2189  | 4,6052  | 5,9915  | 7,3778  | 7,8241  | 9,2104  | 13,8150 |
| 3   | 0,1148 | 0,1848 | 0,2158 | 0,3518 | 0,5844 | 1,0052 | 2,3660 | 4,6416  | 6,2514  | 7,8147  | 9,3484  | 9,8374  | 11,3449 | 16,2660 |
| 4   | 0,2971 | 0,4294 | 0,4844 | 0,7107 | 1,0636 | 1,6488 | 3,3567 | 5,9886  | 7,7794  | 9,4877  | 11,1433 | 11,6678 | 13,2767 | 18,4662 |
| 5   | 0,5543 | 0,7519 | 0,8312 | 1,1455 | 1,6103 | 2,3425 | 4,3515 | 7,2893  | 9,2363  | 11,0705 | 12,8325 | 13,3882 | 15,0863 | 20,5147 |
| 6   | 0,8721 | 1,1344 | 1,2373 | 1,6354 | 2,2041 | 3,0701 | 5,3481 | 8,5581  | 10,6446 | 12,5916 | 14,4494 | 15,0332 | 16,8119 | 22,4575 |
| 7   | 1,2390 | 1,5643 | 1,6899 | 2,1673 | 2,8331 | 3,8223 | 6,3458 | 9,8032  | 12,0170 | 14,0671 | 16,0128 | 16,6224 | 18,4753 | 24,3213 |
| 8   | 1,6465 | 2,0325 | 2,1797 | 2,7326 | 3,4895 | 4,5936 | 7,3441 | 11,0301 | 13,3616 | 15,5073 | 17,5345 | 18,1682 | 20,0902 | 26,1239 |
| 9   | 2,0879 | 2,5324 | 2,7004 | 3,3251 | 4,1682 | 5,3801 | 8,3428 | 12,2421 | 14,6837 | 16,9190 | 19,0228 | 19,6790 | 21,6660 | 27,8767 |
| 10  | 2,5582 | 3,0591 | 3,2470 | 3,9403 | 4,8652 | 6,1791 | 9,3418 | 13,4420 | 15,9872 | 18,3070 | 20,4832 | 21,1608 | 23,2093 | 29,5879 |