

Diseño de experimentos

Objetivos

Presentar los métodos de análisis más comunes para diseños experimentales sencillos.

Discutir condiciones de cada diseño y la interpretación de resultados.

Introducir la metodología de modelos lineales (Análisis de la varianza).

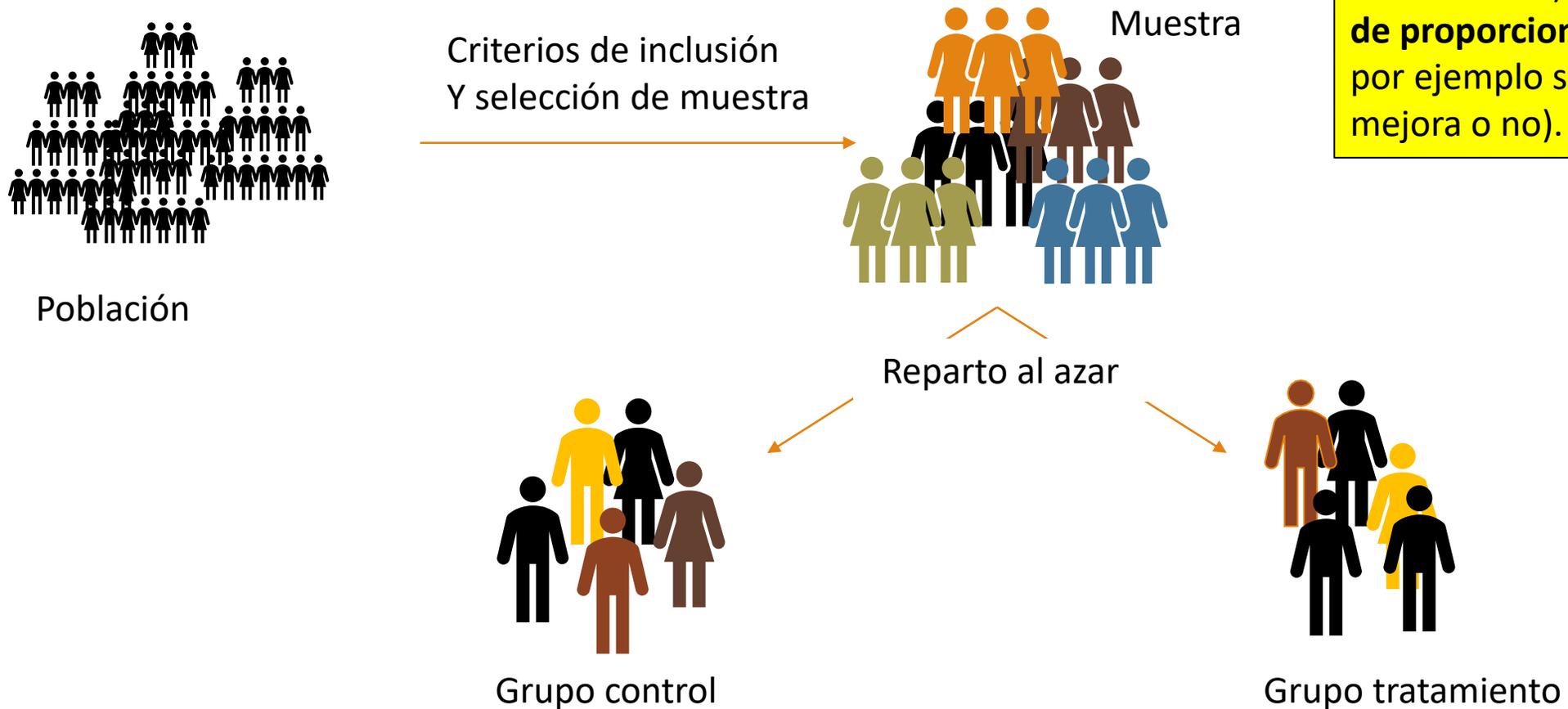
Discutir la interpretación de los intervalos de confianza resultantes y evaluar si el tamaño muestral es suficiente para generalizar los resultados.



Diseño básico

Un grupo control y un tratamiento

El efecto del tratamiento se calcula estimado la **diferencia de medias** entre los dos grupos (si medimos variables cuantitativas) o la **diferencia de proporciones** (si medimos por ejemplo si un paciente mejora o no).



Población

Criterios de inclusión
Y selección de muestra

Muestra

Reparto al azar

Grupo control

Grupo tratamiento

Diseño básico

Un grupo control y un tratamiento

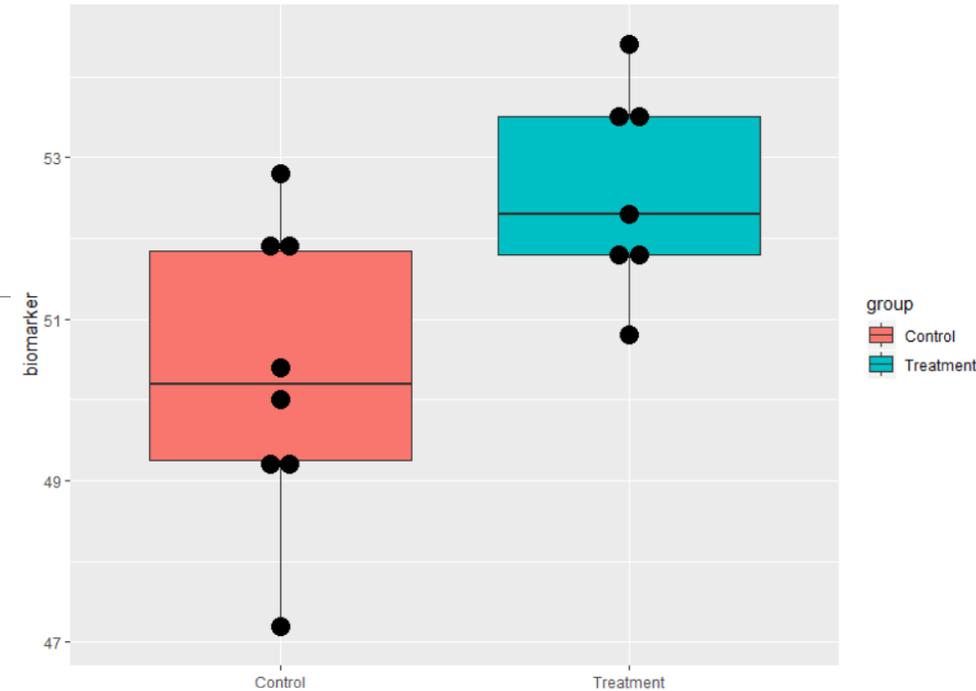
```
> dt
  group biomarker
1 Control    50.4
2 Control    51.8
3 Control    50.0
4 Control    49.1
5 Control    52.8
6 Control    47.2
7 Control    52.0
8 Control    49.3
9 Treatment  54.4
10 Treatment 53.5
11 Treatment 51.7
12 Treatment 53.5
13 Treatment 50.8
14 Treatment 51.9
15 Treatment 52.3
```

	Control N=8	Treatment N=7	p.overall
biomarker	50.3 (1.83)	52.6 (1.26)	0.015

```
> dt %>% t.test(biomarker~group,data=.)
```

Welch Two Sample t-test

```
data: biomarker by group
t = -2.8123, df = 12.376, p-value = 0.01528
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -4.0062819 -0.5151467
sample estimates:
 mean in group Control mean in group Treatment
          50.32500          52.58571
```



Diseño básico

Un grupo control y un tratamiento

```
> dt %>% t.test(biomarker~group,data=.)
```

```
Welch Two Sample t-test
```

```
data: biomarker by group
t = -2.8123, df = 12.376, p-value = 0.01528
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -4.0062819 -0.5151467
sample estimates:
 mean in group Control mean in group Treatment
                50.32500                52.58571
```

El IC de confianza de la diferencia de medias poblacionales entre los grupos es (-4.0, -0.5). Este resultado indica que podemos concluir que el tratamiento produce un incremento en la media del valor del biomarcador de entre 0.5 y 4 unidades.

Dado que estamos trabajando con muy pocos individuos, este intervalo es amplio y deberíamos aumentar el tamaño muestral para conseguir una estimación más precisa.

Deberíamos especificar qué diferencia es relevante desde el punto de vista del efecto del tratamiento y su consideración para aplicación en pacientes. Si el efecto mínimo requerido fuera de 5 unidades, entonces este tratamiento no cumpliría con dicho requisito.

Diseño básico

Un grupo control y un tratamiento

```
> dt %>% t.test(biomarker~group,data=.)
```

```
Welch Two Sample t-test
```

```
data: biomarker by group
t = -2.8123, df = 12.376, p-value = 0.01528
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -4.0062819 -0.5151467
sample estimates:
 mean in group Control mean in group Treatment
                50.32500                52.58571
```

El IC de confianza de la diferencia de medias poblacionales entre los grupos es (-4.0, -0.5). Este resultado indica que podemos concluir que el tratamiento produce un incremento en la media del valor del biomarcador de entre 0.5 y 4 unidades.

Dado que estamos trabajando con muy pocos individuos, este intervalo es amplio y deberíamos aumentar el tamaño muestral para conseguir una estimación más precisa.

Deberíamos especificar qué diferencia es relevante desde el punto de vista del efecto del tratamiento y su consideración para aplicación en pacientes. Si el efecto mínimo requerido fuera de 5 unidades, entonces este tratamiento no cumpliría con dicho requisito.

Diseño básico

Un grupo control y un tratamiento

```
> dt %>% t.test(biomarker~group,data=.)
```

Welch Two Sample t-test

data: biomarker by group

t = -2.8123, df = 12.376, p-value = 0.01528

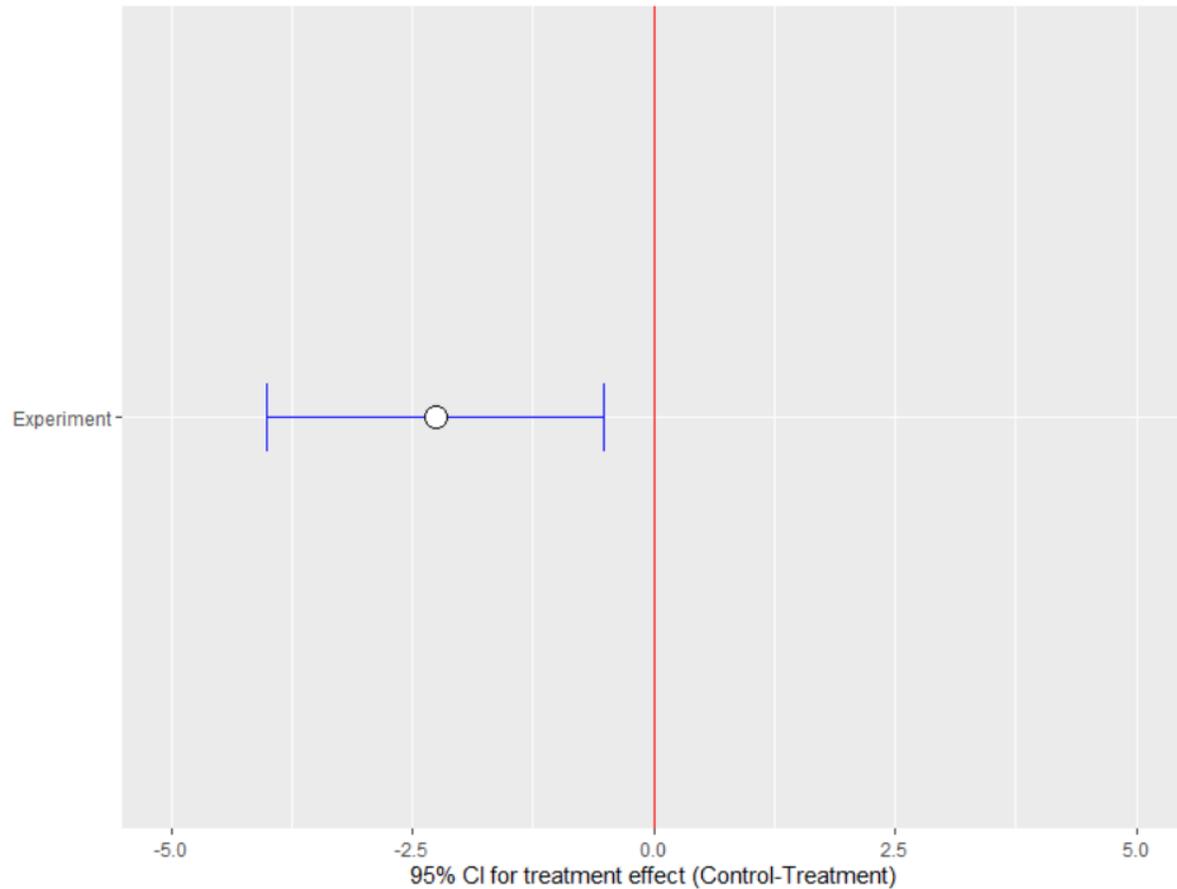
alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-4.0062819 -0.5151467

sample estimates:

mean in group Control	mean in group Treatment
50.32500	52.58571

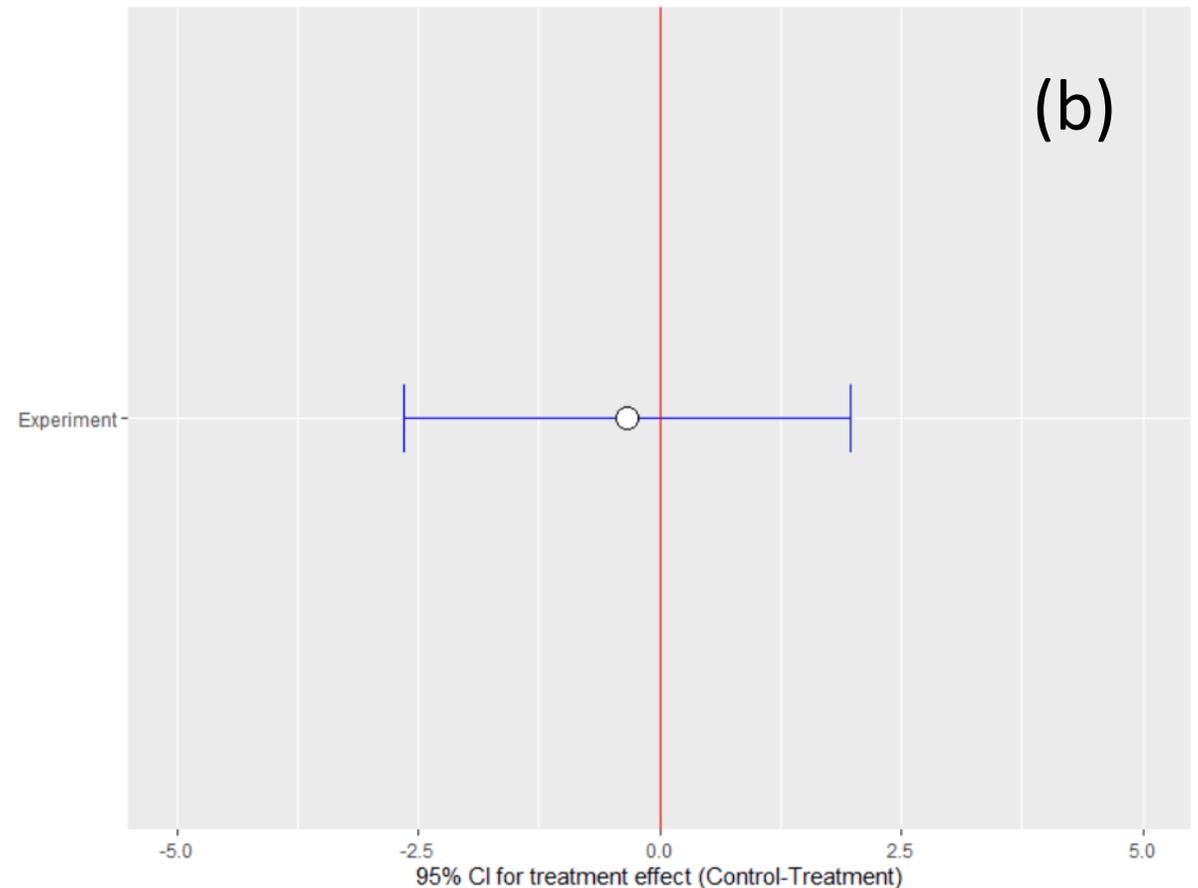
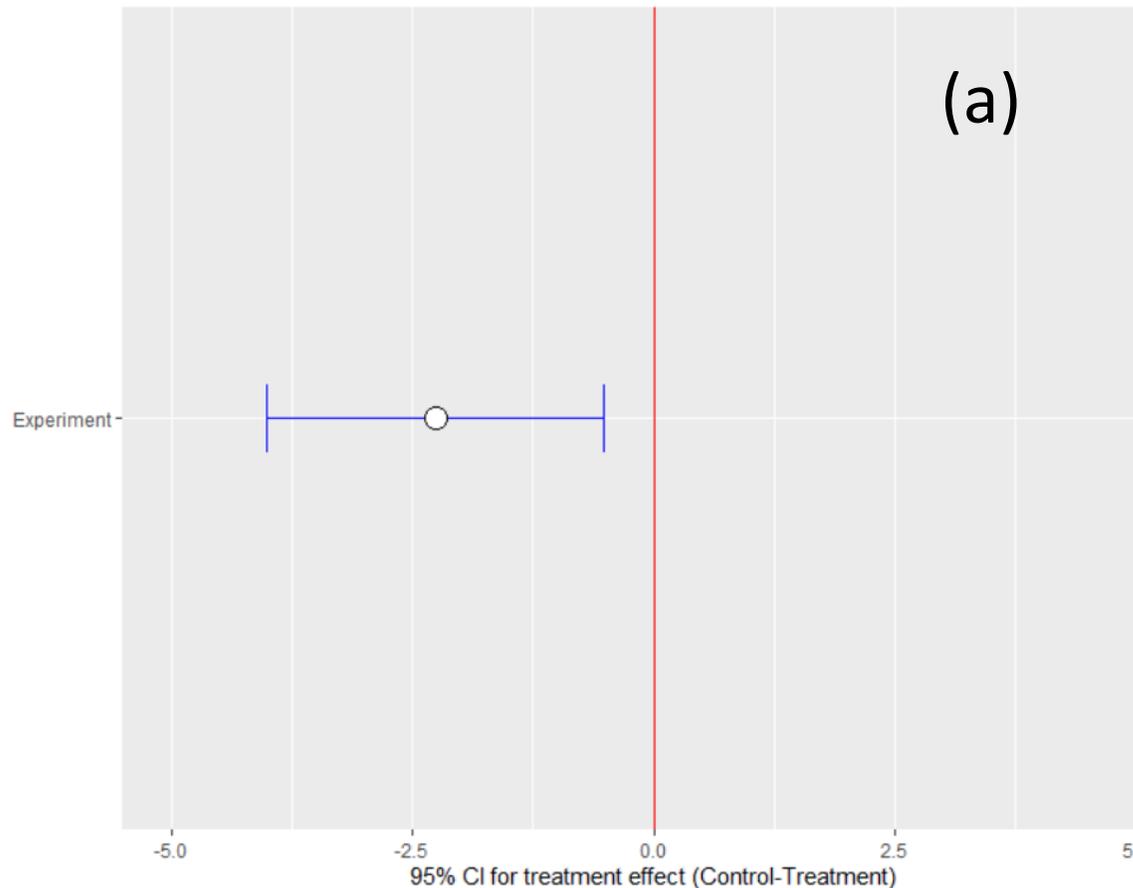


Diseño básico

Un grupo control y un tratamiento

Resultado de dos experimentos. Se puede comprobar que en el caso (a) la conclusión sería que existe un efecto del tratamiento, mientras que en el caso (b) no.

Trabajar con muestras pequeñas conlleva incertidumbre en las conclusiones.



En la base de datos birthwt (librería MASS) se recoge el peso de recién nacidos y los datos de las madres. Queremos verificar si el peso (**bwt** en gramos) depende de si la madre es fumadora o no (smoke, 1=Fumadora, 0=No fumadora).

```
> library(MASS)
> head(birthwt)
  low age lwt race smoke ptl ht ui ftv bwt
85   0  19 182   2     0   0  0  1   0 2523
86   0  33 155   3     0   0  0  0   3 2551
87   0  20 105   1     1   0  0  0   1 2557
88   0  21 108   1     1   0  0  1   2 2594
89   0  18 107   1     1   0  0  1   0 2600
91   0  21 124   3     0   0  0  0   0 2622

> birthwt %>% t.test(bwt~smoke,data=.)

Welch Two Sample t-test

data:  bwt by smoke
t = 2.7299, df = 170.1, p-value = 0.007003
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 78.57486 488.97860
sample estimates:
mean in group 0 mean in group 1
 3055.696      2771.919
```

El mismo procedimiento puede utilizarse para comparar dos grupos (cohortes). Por ejemplo, en esta base de datos tenemos información acerca de si la madre es fumadora.

Podemos verificar si en función del hábito de fumar, sin tener en cuenta otros posibles factores, el peso del recién nacido es menor.

El IC (No fumadoras vs. Fumadoras) es (78.6, 488.9). Recordad que el peso está medido en gramos.

Por lo tanto, podemos concluir que el hecho de fumar (group 1) produce una disminución del peso medio de los recién nacidos.

Dado que el peso está en gramos, si lo pasamos a kg. el IC es de (0.08, 0.49). Es decir, la disminución de peso sería inferior al medio kg.

Estudio experimental (Control vs tratamiento)

Datos cualitativos

La evaluación del efecto de un tratamiento puede hacerse también a partir de la proporción de mejoras. Formalmente, es lo que se conoce como estudio de cohortes, donde disponemos de grupos definidos, ya sea de manera natural (hombres/mujeres) o por intervención (control/tratamiento).

La estimación del efecto del tratamiento se realiza mediante el IC de la diferencia de probabilidades o calculando el riesgo relativo.

En el tema de estimación de parámetros se han discutido algunos ejemplos. A continuación mostraremos algunos más para consolidar conceptos.

Estudio experimental (Control vs tratamiento)

Datos cualitativos

```
df <- matrix(c(22,25,25,17),ncol=2,byrow=T)
colnames(df)<- c('Sanos','Enfermos')
rownames(df)<- c('Control','Tratamiento')
addmargins(df,2)
```

	Sanos	Enfermos	Sum
Control	22	25	47
Tratamiento	25	17	42

```
> addmargins(prop.table(df,1)*100,2) %>% round(2)
```

	Sanos	Enfermos	Sum
Control	46.81	53.19	100
Tratamiento	59.52	40.48	100

```
> riskratio(df)$measure %>% round(2)
```

	NA	estimate	lower	upper
risk ratio with 95% C.I.				
Control	1.00	NA	NA	NA
Tratamiento	0.76	0.48	1.2	

El IC del RR (0.48, 1.2) no permite concluir que el tratamiento reduzca la proporción de enfermos.

Estudio experimental (Control vs tratamiento)

Datos cualitativos (a partir de un data.frame)

Tenemos un data.frame de nombre *experim*, que contiene los resultados individuales de la aparición de una enfermedad en dos grupos experimentales. Los 15 primeros casos son:

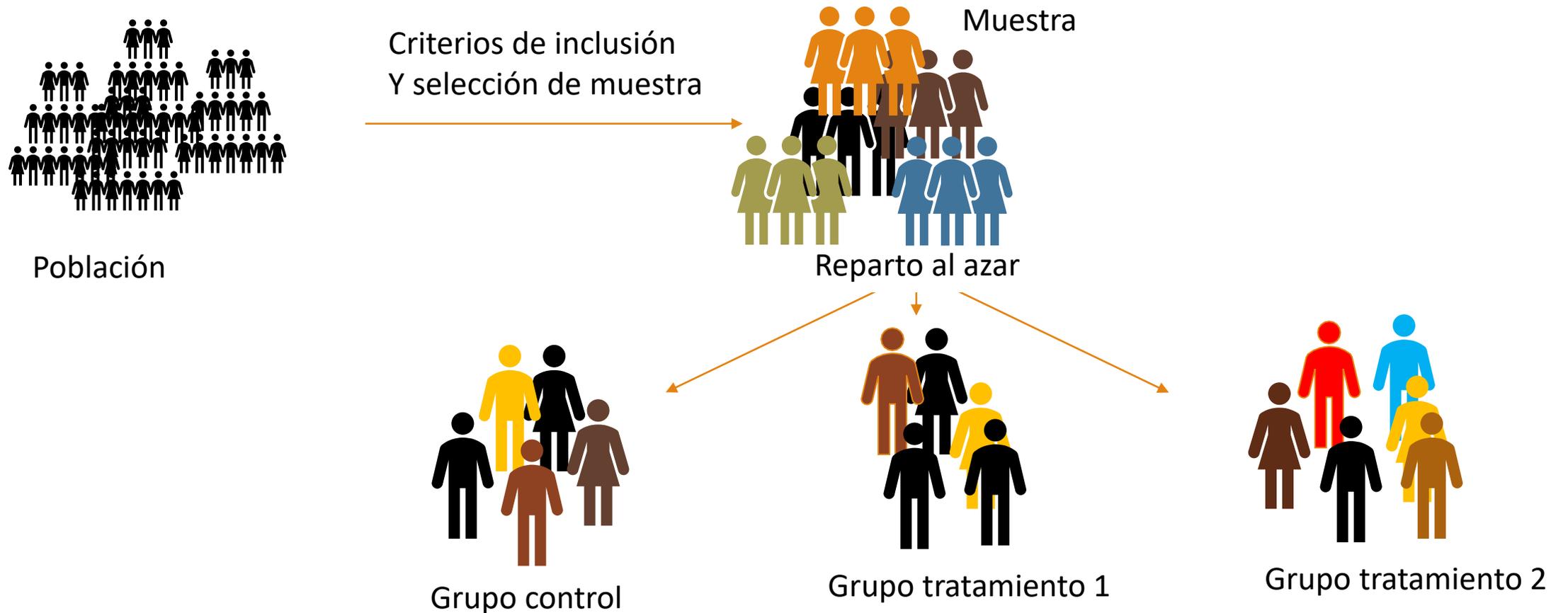
	Group	Outcome
1	Control	Sanos
2	Control	Sanos
3	Control	Sanos
4	Control	Sanos
5	Control	Sanos
6	Control	Sanos
7	Control	Sanos
8	Control	Sanos
9	Control	Sanos
10	Control	Sanos
11	Control	Sanos
12	Control	Sanos
13	Control	Enfermos
14	Control	Enfermos
15	Control	Enfermos

```
t <- with(experim, table(Group, Outcome))
addmargins(prop.table(t, 1), 2) %>% round(2)
riskratio(t)$measure %>% round(2)
```

Group	Outcome		Sum
	Sanos	Enfermos	
Control	0.27	0.73	1.00
Tratamiento	0.32	0.68	1.00

Group	risk ratio with 95% C.I.		
	estimate	lower	upper
Control	1.00	NA	NA
Tratamiento	0.94	0.73	1.2

Un control (grupo de referencia) y varios tratamientos



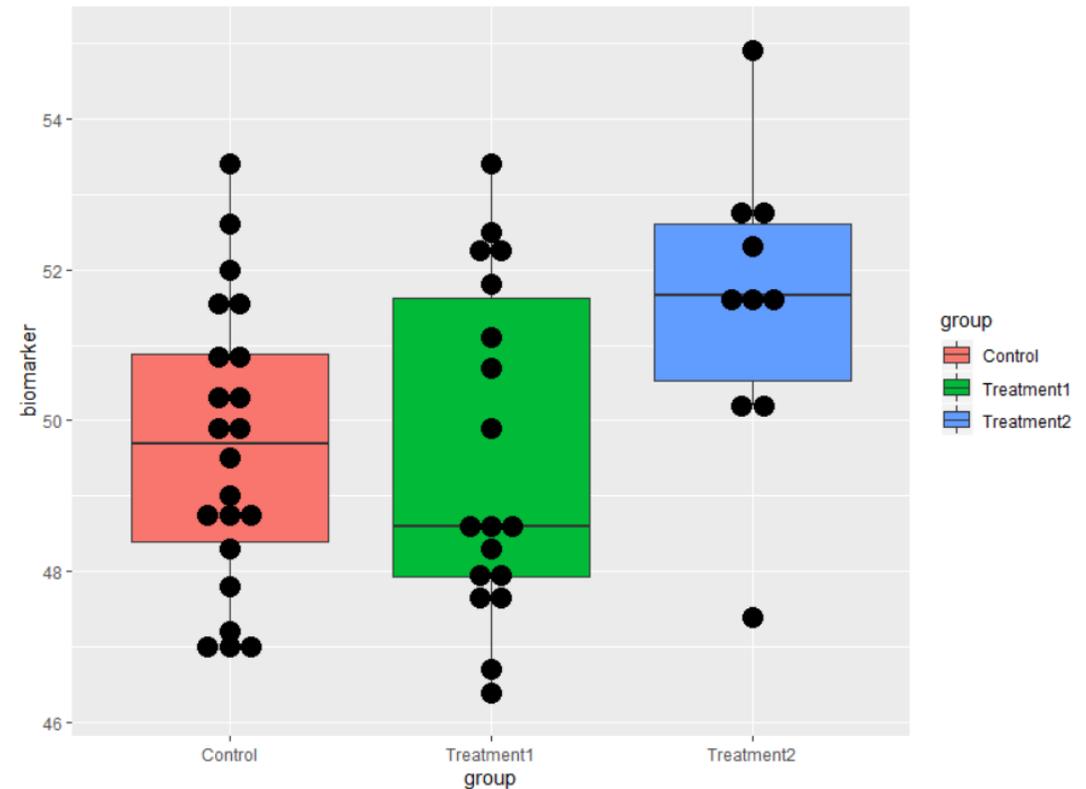
Un control (grupo de referencia) y varios tratamientos

¿Cómo analizar los resultados de un experimento donde tenemos un grupo control y más de un tratamiento?

En la gráfica se representa un resultado de valores de un biomarcador en tres grupos (Control, Treatment1, Treatment2).

Los datos tienen esta forma:

```
> dt %>% head()
  group biomarker
1: Treatment1    48.6
2: Treatment1    48.6
3:   Control    46.9
4: Treatment1    49.9
5: Treatment2    52.7
6:   Control    49.5
```



Un control (grupo de referencia) y varios tratamientos

Calculamos la media, desviación estándar y tamaño muestral para cada grupo.

Se trata de determinar si la media varía en función del grupo y estimar el efecto de cada tratamiento respecto al control.

No es correcto utilizar el IC de la diferencia comparando cada tratamiento con el control.

Debemos analizar los datos utilizando la técnica de análisis de la varianza (ANOVA).

```
dt %>% group_by(group) %>%  
  summarise(Mean=round(mean(biomarker),2),  
            SD=round(sd(biomarker),2),  
            N=n()) %>%  
  as.data.frame
```

	group	Mean	SD	N
	Control	49.65	1.88	22
	Treatment1	49.57	2.19	18
	Treatment2	51.53	1.99	10

Análisis de la varianza

Un factor

Modelo lineal (un factor fijo):

- El factor fijo es una condición experimental, p.e. Control, Trat1, Trat2.
- En ausencia de efecto debido al tratamiento, la variable respuesta sigue una distribución $N(\mu, \sigma)$.
- Si existe un efecto tratamiento, la distribución para el tratamiento i es $N(\mu + \alpha_i, \sigma)$
- Cada grupo experimental tiene n_i individuos
- La observación j en el grupo i se expresa como:

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij} \quad \epsilon_{ij} \rightarrow N(0, \sigma)$$

- Se trata de estimar los efectos α_i y determinar si son distintos de cero.

G_1	G_2	G_k	
y_{11}	y_{21}	y_{k1}	
y_{12}	y_{22}	y_{k2}	
\vdots	\vdots	\ddots	\ddots		
y_{1n_1}	y_{2n_2}	y_{kn_k}	
$\bar{y}_1.$	$\bar{y}_2.$	$\bar{y}_k.$	$\bar{y}_{..}$

Análisis de la varianza

Un factor

Análisis `res <- lm(biomarker~group,data=dt)`
`anova(res)`

La función **lm** (linear model) ajusta un modelo lineal a los datos, estimando los efectos de cada tratamiento y evaluando su significación.

Resultados

Analysis of Variance Table

```
Response: biomarker
          Df Sum Sq Mean Sq F value Pr(>F)
group      2  29.398  14.6988   3.6096 0.03481 *
Residuals 47 191.392   4.0722
```

El análisis de la varianza evalúa si se puede aceptar que los efectos son iguales a cero.

El p-valor (en este caso 0.035) es la probabilidad de haber obtenido los resultados que tenemos o un resultado aún más alejad del esperado si los efectos son cero.

Un valor bajo del p-valor nos indica que los resultados no son del todo compatibles con la situación de efectos igual a cero.

Análisis de la varianza

Un factor

Análisis `confint(res)`
`TukeyHSD(aov(res))`

La función **confint** (linear model) proporciona los IC de los efectos tomando el primer grupo como referencia.

La función **TukeyHSD** proporciona IC para cada par de comparaciones.

El tratamiento 1 no aumenta la media respecto al control IC: (-1.63, 1.47).

El tratamiento aumenta la media IC: (0.02,3.74) aunque deberíamos disponer de más datos.

Resultados

	2.5 %	97.5 %
(Intercept)	48.7844868	50.515513
groupTreatment1	-1.3680087	1.212453
groupTreatment2	0.3317229	3.428277

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = res)

\$group

	diff	lwr	upr	p adj
Treatment1-Control	-0.07777778	-1.62992418	1.474369	0.9919261
Treatment2-Control	1.88000000	0.01742432	3.742576	0.0474500
Treatment2-Treatment1	1.95777778	0.03161475	3.883941	0.0456017

Análisis de la varianza

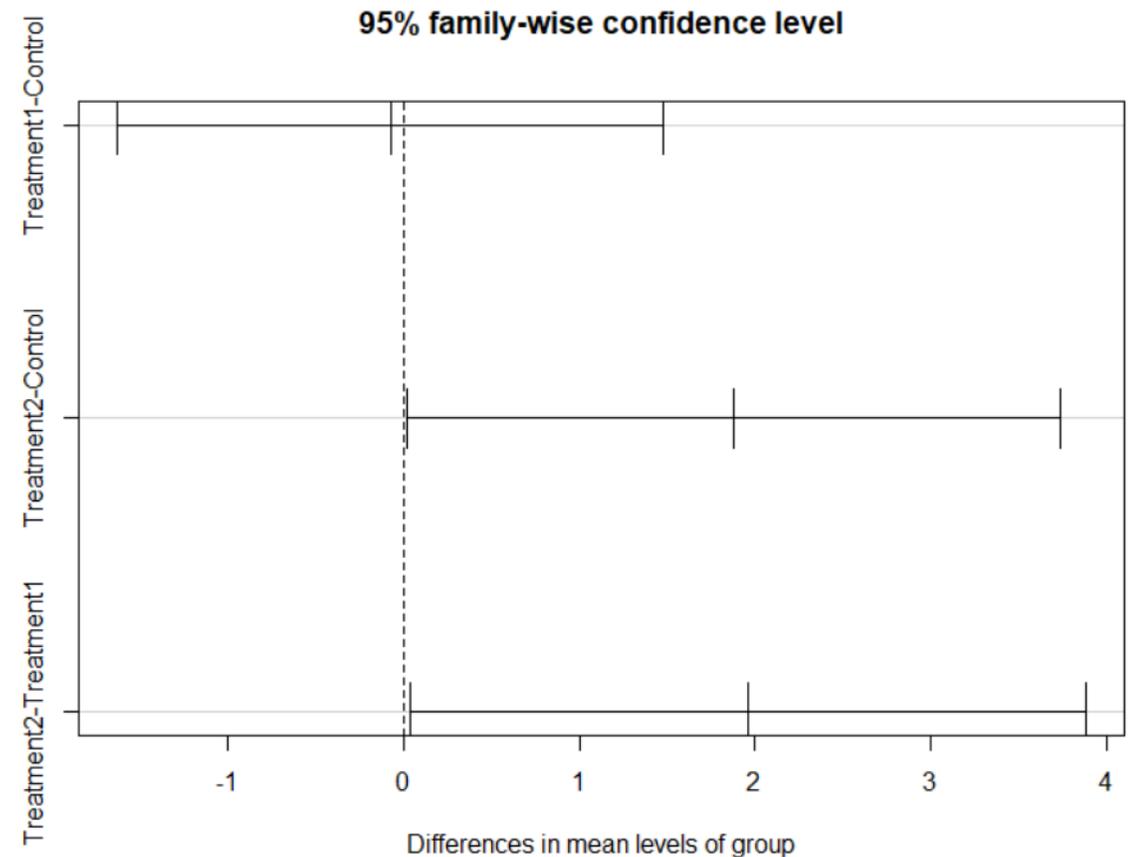
Un factor

Análisis `TukeyHSD(aov(res)) %>% plot()`

Mediante **plot** obtenemos la gráfica de los intervalos calculados por la función **TukeyHSD**.

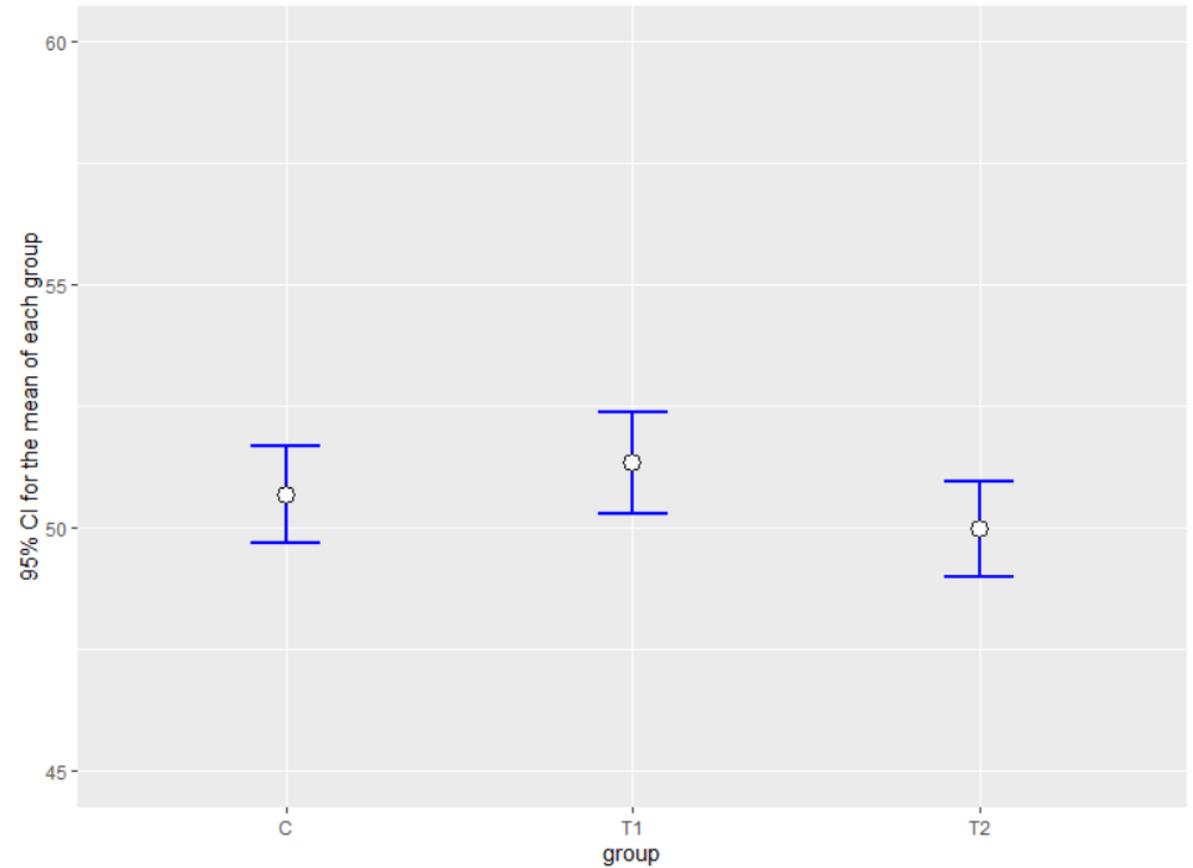
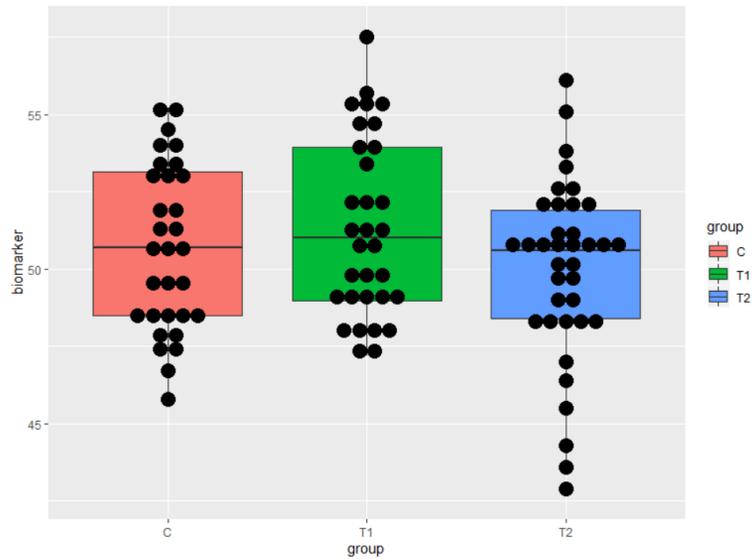
El IC T1-C indica que el T1 se comporta de manera similar que el control.

El IC T2-C indica que T2 da valores medios superiores, aunque puede que la diferencia sea pequeña. Lo mismo para T2-T1.



Ejemplos de interpretación (1)

	group	Mean	SD	N
1	C	50.69	2.69	31
2	T1	51.34	2.89	32
3	T2	49.98	2.93	37



Ejemplos de interpretación (1)

Analysis of Variance Table

Response: biomarker

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
group	2	31.97	15.9848	1.9725	0.1446
Residuals	97	786.05	8.1037		

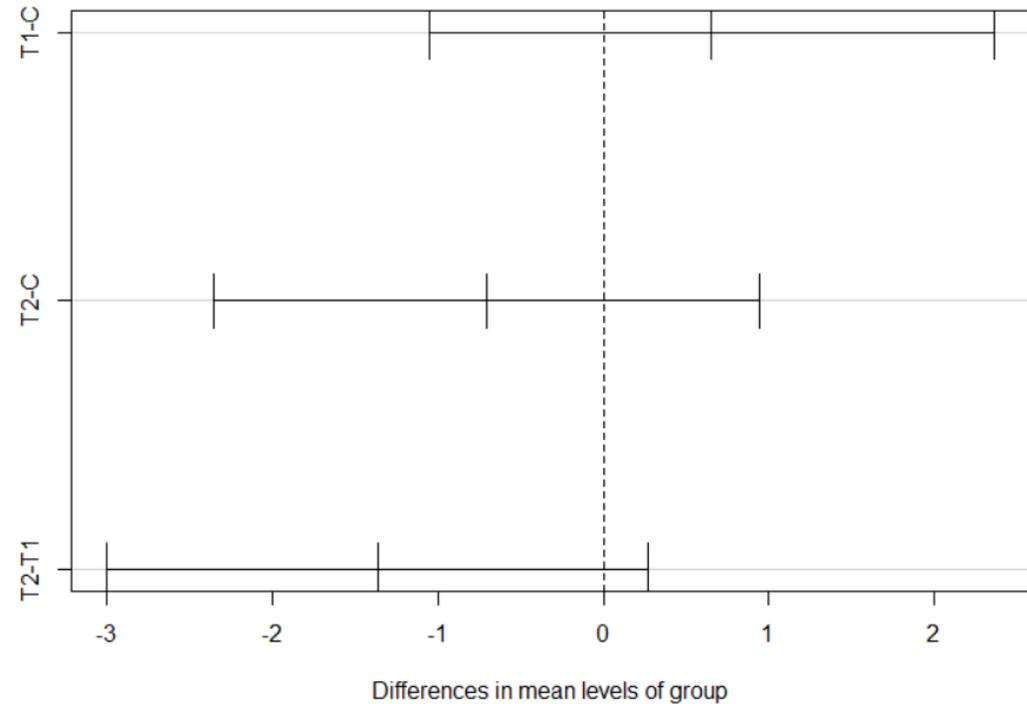
	2.5 %	97.5 %
(Intercept)	49.6723455	51.7018481
groupT1	-0.7702912	2.0773476
groupT2	-2.0843851	0.6669483

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = res)

\$group	diff	lwr	upr	p adj
T1-C	0.6535282	-1.054017	2.3610737	0.6347014
T2-C	-0.7087184	-2.358516	0.9410789	0.5645314
T2-T1	-1.3622466	-2.997957	0.2734642	0.1220361

95% family-wise confidence level



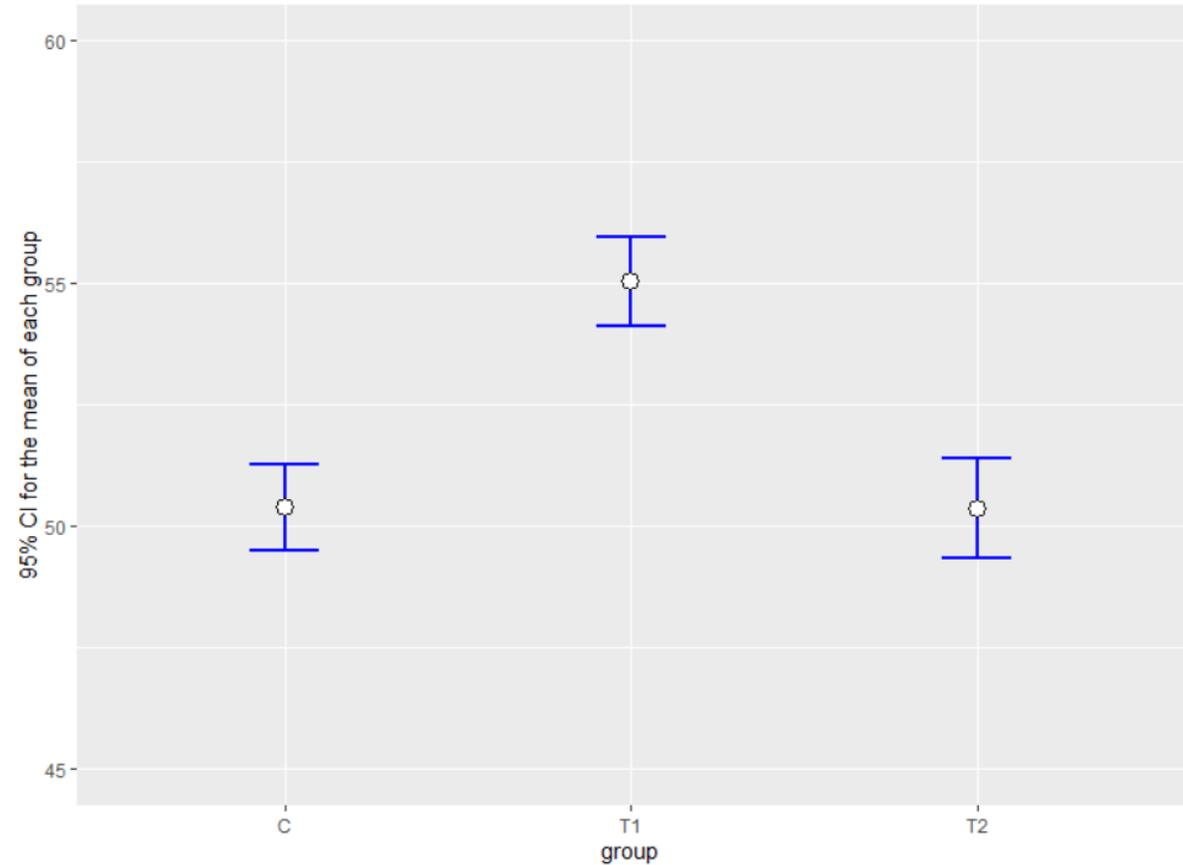
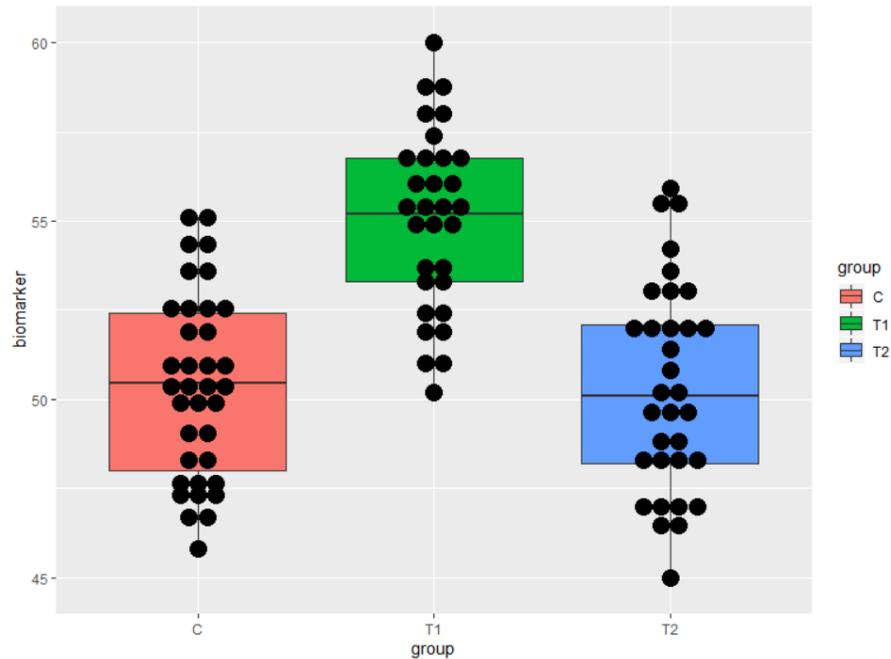
La tabla de ANOVA indica ($p\text{-valor}=0.14$) que los datos son compatibles con un efecto nulo en todos los grupos.

La estimación de los efectos produce IC que incluyen el efecto 0 en todos los grupos.

Los IC de las diferencias entre grupos incluyen el valor 0.

Ejemplos de interpretación (2)

	group	Mean	SD	N
1	C	50.39	2.58	36
2	T1	55.05	2.52	31
3	T2	50.37	2.92	33



Ejemplos de interpretación (2)

Analysis of Variance Table

Response: biomarker

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
group	2	466.88	233.439	32.532	1.544e-11 ***
Residuals	97	696.04	7.176		

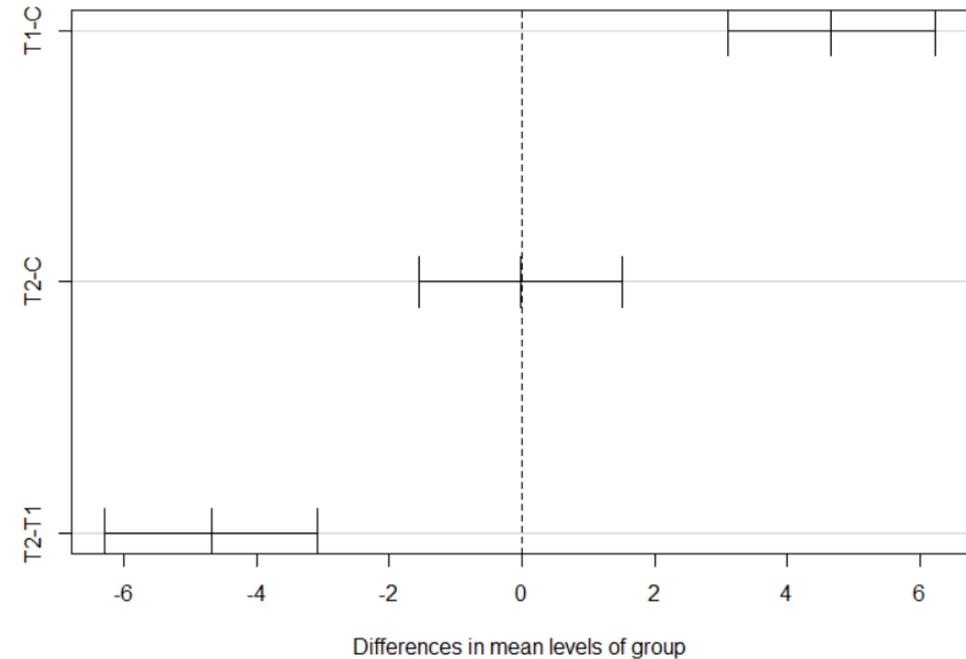
	2.5 %	97.5 %
(Intercept)	49.502793	51.274985
groupT1	3.360045	5.965403
groupT2	-1.300485	1.262101

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = res)

\$group	diff	lwr	upr	p adj
T1-C	4.66272401	3.100459	6.224989	0.0000000
T2-C	-0.01919192	-1.555809	1.517425	0.9995129
T2-T1	-4.68191593	-6.276699	-3.087133	0.0000000

95% family-wise confidence level



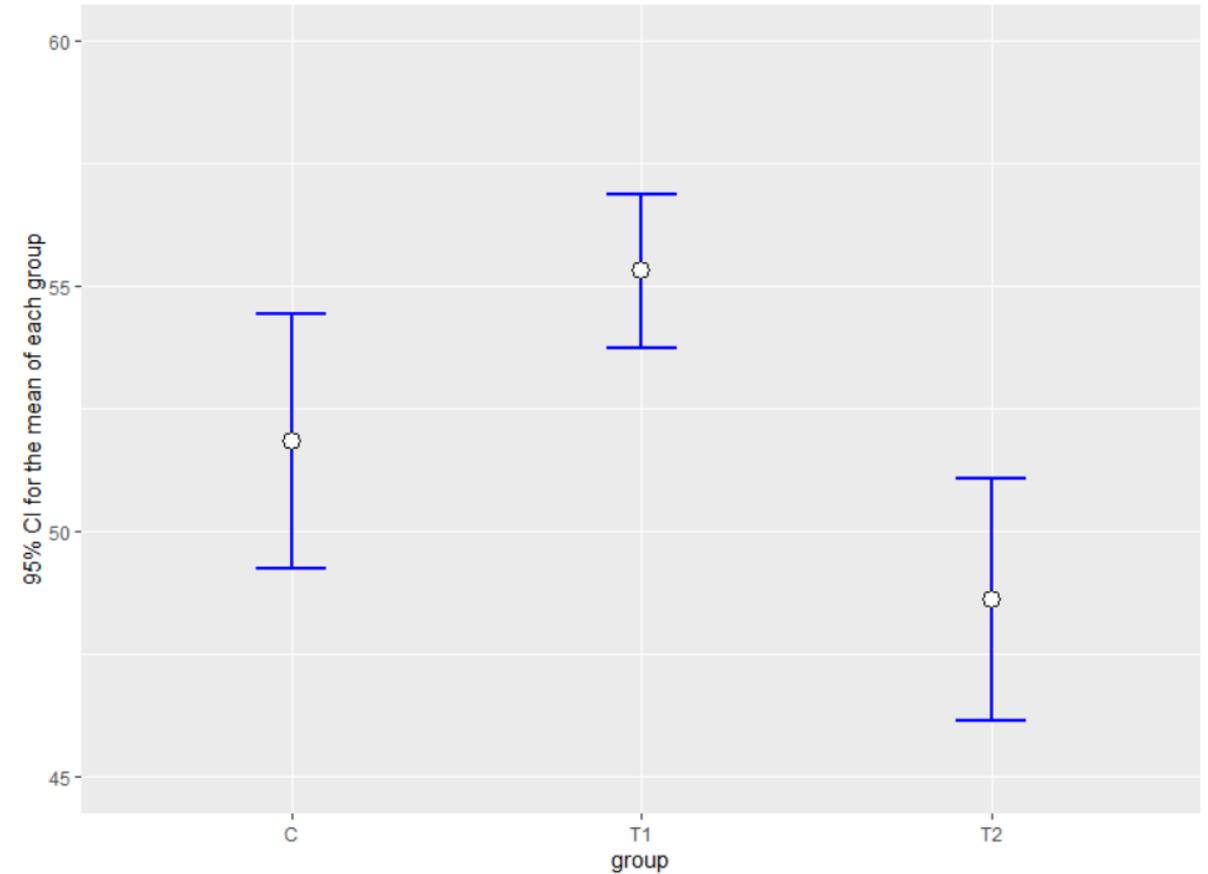
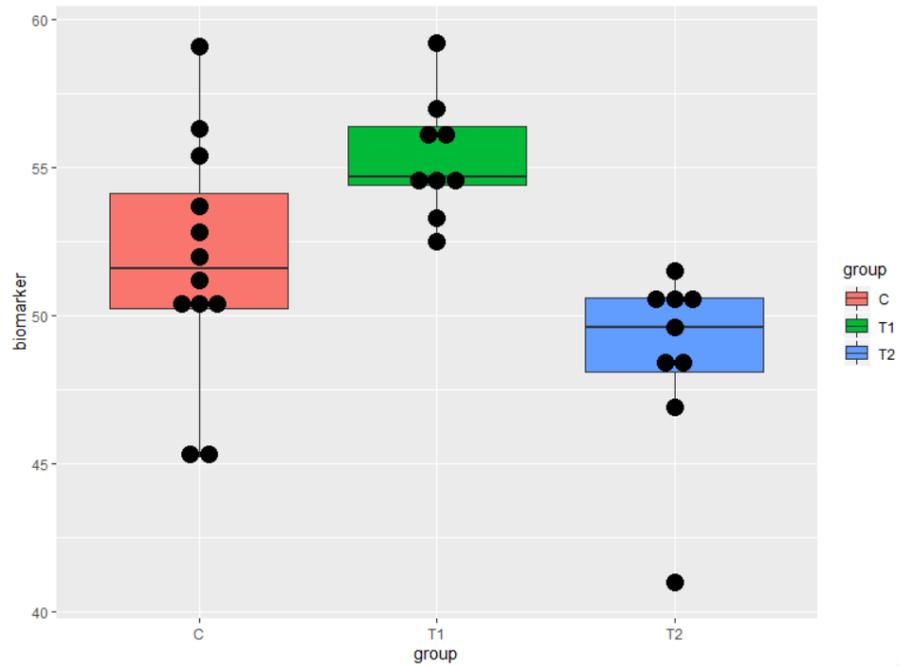
La tabla de ANOVA indica (***p-valor*** muy pequeño) que los datos **no son compatibles** con un efecto nulo en todos los grupos.

La estimación de los efectos produce IC que indican un efecto importante del T1.

Los IC de las diferencias entre indican un efecto de T1 respecto al C (IC: 3.1, 6.2), un efecto cero de T2 (IC -1.6,1.5), y una diferencia en el efecto de T1 respecto al efecto de T2 (IC 3.1,6.3).

Ejemplos de interpretación (3)

group	Mean	SD	N
C	51.85	4.07	12
T1	55.32	2.03	9
T2	48.61	3.20	9



Ejemplos de interpretación (3)

Analysis of Variance Table

Response: biomarker

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
group	2	202.77	101.387	9.2029	0.0008962	***
Residuals	27	297.45	11.017			

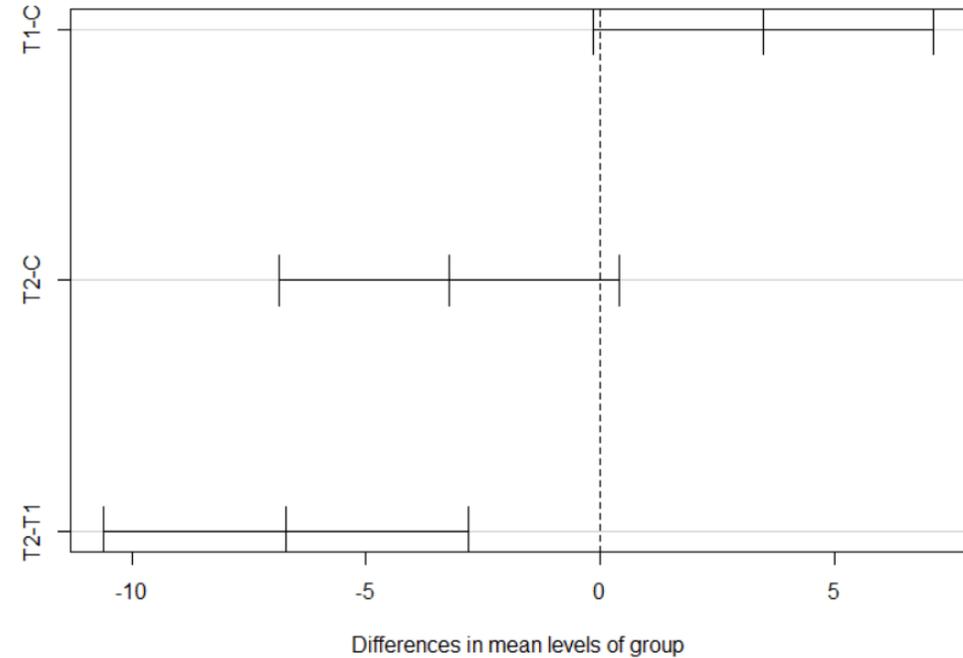
	2.5 %	97.5 %
(Intercept)	49.8840195	53.815981
groupT1	0.4691374	6.475307
groupT2	-6.2419737	-0.235804

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = res)

\$group	diff	lwr	upr	p adj
T1-C	3.472222	-0.1566846	7.1011290	0.0627568
T2-C	-3.238889	-6.8677957	0.3900179	0.0870194
T2-T1	-6.711111	-10.5905757	-2.8316466	0.0005835

95% family-wise confidence level



La tabla de ANOVA indica ($p\text{-valor}=0.0009$) que los datos **no son compatibles** con un efecto nulo en todos los grupos.

Los IC de la comparación de grupos no permiten concluir un efecto de T1 o T2 respecto al C, aunque cuando se comparan T1 y T2 el IC (2.8,10.6) permite concluir que el efecto de T1 es mayor al de T2.

Sin embargo, las limitaciones del tamaño muestral determinan que estos resultados no sean muy concluyentes.

Diseño con dos factores fijos

Queremos determinar si la producción de un nutriente esencial en biorreactores depende de la temperatura y del sustrato de crecimiento de las bacterias que lo producen.

Para ello, diseñamos un experimento donde se disponen tres sustratos (A, B, C) y para cada sustrato se verifica la producción a dos temperaturas distintas (30 i 35°C).

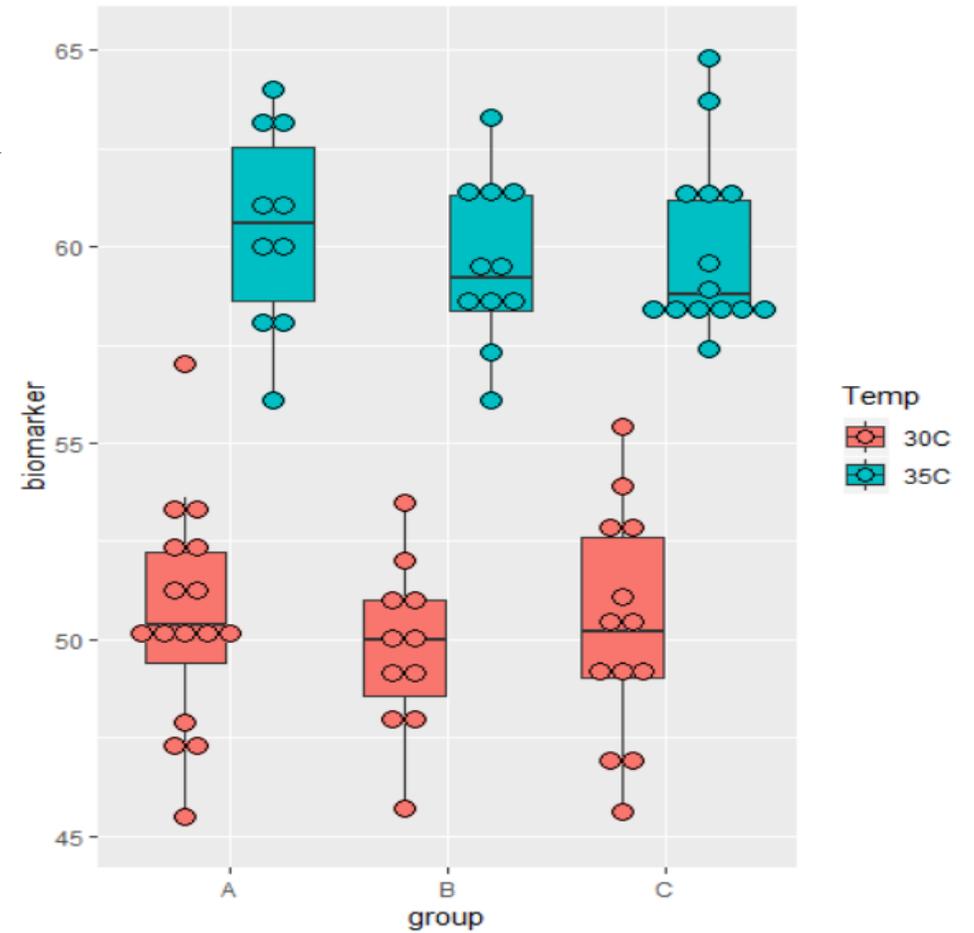
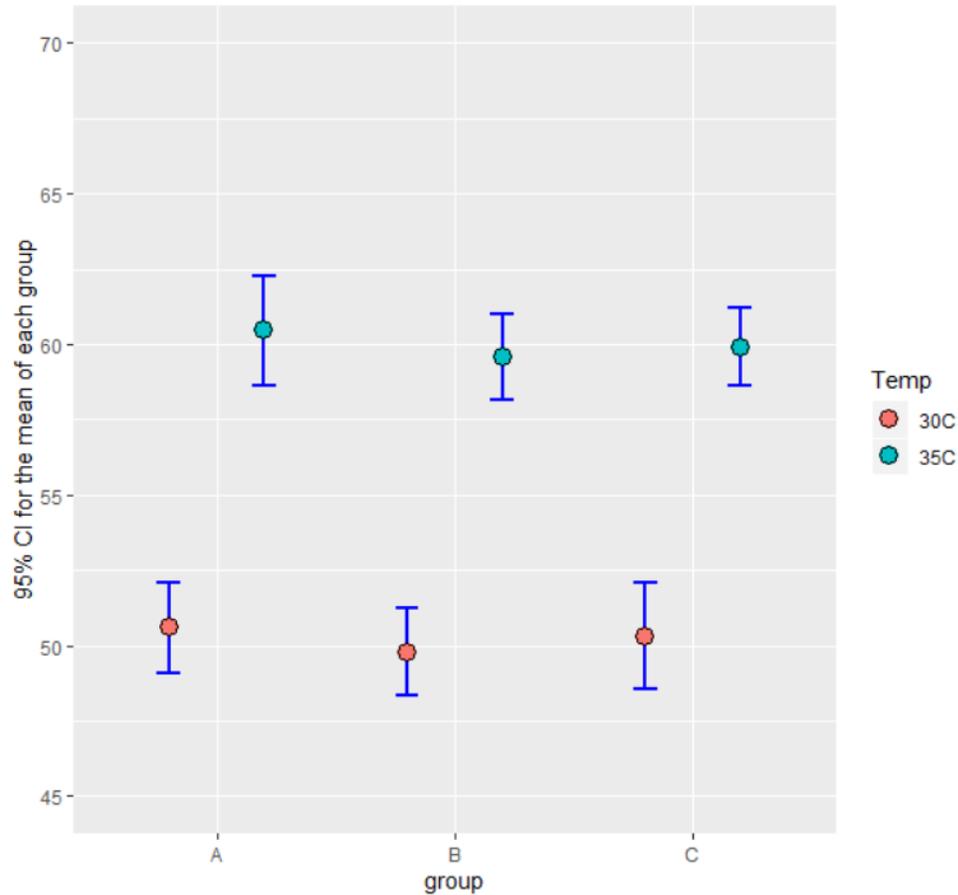
En este caso, tenemos dos factores (Sustrato y Temperatura) y queremos verificar si la producción media depende de ellos.

La tabla indica la disposición del experimento. El factor sustrato tiene tres niveles (A, B, C) y el factor temperatura tiene dos niveles (30 i 35°C).

		A	B	C
Temperatura	30 °C
	
	35 °C
	

group	Temp	Mean	SD	N
A	30C	50.60	2.82	16
A	35C	60.46	2.56	10
B	30C	49.77	2.15	11
B	35C	59.59	2.12	11
C	30C	50.31	2.91	13
C	35C	59.91	2.24	14

Diseño con dos factores fijos



Parece claro que la producción del nutriente depende de la temperatura y que para cada temperatura no hay grandes diferencias entre substratos.

Diseño con dos factores fijos : análisis

```
res <- lm(biomarker~group*Temp,data=dt)
anova(res)
```

Analysis of Variance Table

Response: biomarker

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
group	2	11.08	5.54	0.8757	0.4211
Temp	1	1750.30	1750.30	276.6279	<2e-16 ***
group:Temp	2	0.24	0.12	0.0188	0.9814
Residuals	69	436.58	6.33		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Ajustamos el modelo donde biomarker depende del grupo, la temperatura y de su interacción.

El análisis muestra que el efecto del grupo no es significativo ($p=0.42$). El efecto de la temperatura es claro (p muy pequeño) y la interacción no es significativo ($p=0.98$)

Diseño con dos factores fijos : análisis

Estimación de los efectos

```
> coefficients(res)
```

```
(Intercept)          groupB          groupC          Temp35C groupB:Temp35C groupC:Temp35C  
50.60000000    -0.82727273    -0.29230769     9.86000000    -0.04181818    -0.25340659
```

```
> confint(res)
```

```
                2.5 %    97.5 %  
(Intercept)    49.345476  51.854524  
groupB         -2.792734   1.138189  
groupC         -2.166035   1.581420  
Temp35C        7.837141  11.882859  
groupB:Temp35C -2.986369   2.902732  
groupC:Temp35C -3.051200   2.544387
```

La media para el grupo A (que actúa como referencia) y temperatura 30C es 50.6. Para la temperatura 35C sería $50.6+9.86$.

El efecto de la temperatura a 35C es de un aumento medio de 9.86 respecto a la media en 30C (IC 7.8, 11.9) para cualquier grupo.

Para 30C, la media para el Grupo B es $50.6-0.83$ y para 35C sería $50-0.83+9.86$.

Diseño con dos factores fijos IC para diferencia entre grupos

Los IC de comparación de la producción entre los distintos substratos muestran que no hay diferencias.

Para una temperatura dada, tampoco hay diferencias entre substratos.

Para cada substrato, la diferencia entre temperaturas es significativa.

```
> TukeyHSD(aov(res))  
Tukey multiple comparisons of means  
95% family-wise confidence level
```

```
Fit: aov(formula = res)
```

```
$group  
      diff      lwr      upr      p adj  
B-A 0.2895105 -1.455880 2.034901 0.9167756  
C-A 0.8965812 -0.758958 2.552120 0.4015778  
C-B 0.6070707 -1.123442 2.337584 0.6794155
```

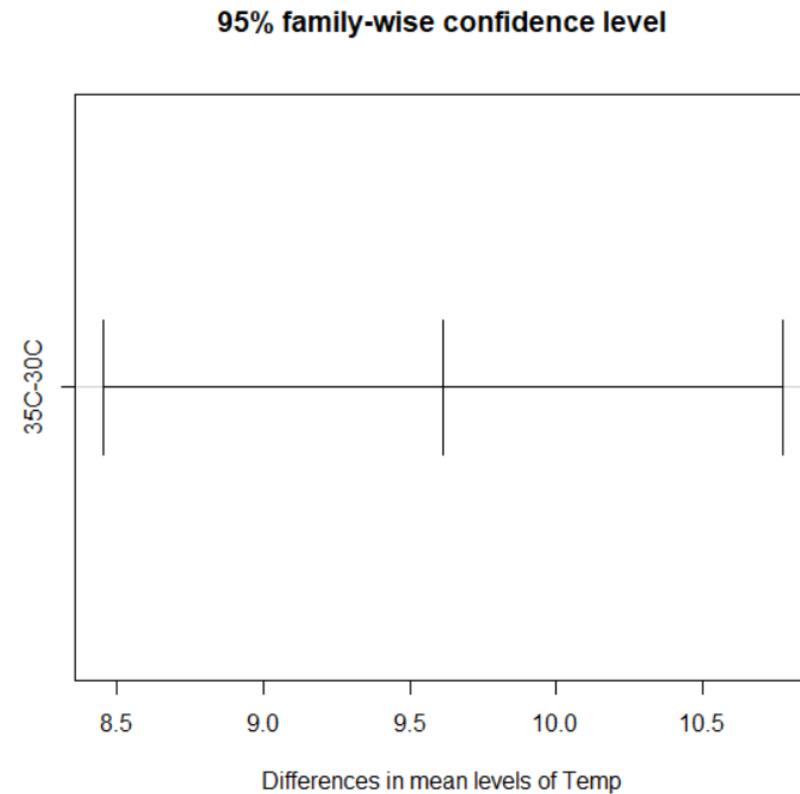
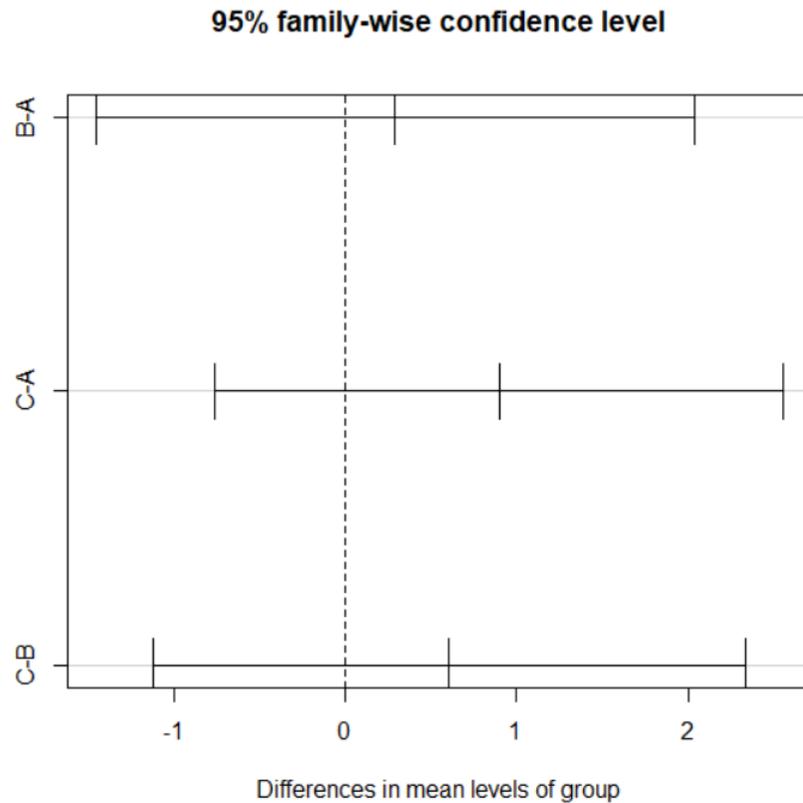
```
$Temp  
      diff      lwr      upr      p adj  
35C-30C 9.612454 8.450991 10.77392      0
```

```
$`group:Temp`  
      diff      lwr      upr      p adj  
B:30C-A:30C -0.8272727 -3.715270 2.060724 0.9589486  
C:30C-A:30C -0.2923077 -3.045512 2.460897 0.9995929  
A:35C-A:30C 9.8600000 6.887664 12.832336 0.0000000  
B:35C-A:30C 8.9909091 6.102912 11.878906 0.0000000  
C:35C-A:30C 9.3142857 6.615877 12.012695 0.0000000  
C:30C-B:30C 0.5349650 -2.485743 3.555673 0.9952479  
A:35C-B:30C 10.6872727 7.465577 13.908968 0.0000000  
B:35C-B:30C 9.8181818 6.674129 12.962235 0.0000000  
C:35C-B:30C 10.1415584 7.170708 13.112409 0.0000000  
A:35C-C:30C 10.1523077 7.050867 13.253748 0.0000000  
B:35C-C:30C 9.2832168 6.262508 12.303925 0.0000000  
C:35C-C:30C 9.6065934 6.766600 12.446587 0.0000000  
B:35C-A:35C -0.8690909 -4.090786 2.352605 0.9682036  
C:35C-A:35C -0.5457143 -3.598616 2.507187 0.9950365  
C:35C-B:35C 0.3233766 -2.647474 3.294227 0.9995402
```

Diseño con dos factores fijos

IC para diferencia entre grupos

```
TukeyHSD(aov(res), 'group') %>% plot()  
TukeyHSD(aov(res), 'Temp') %>% plot()
```



Ejemplo

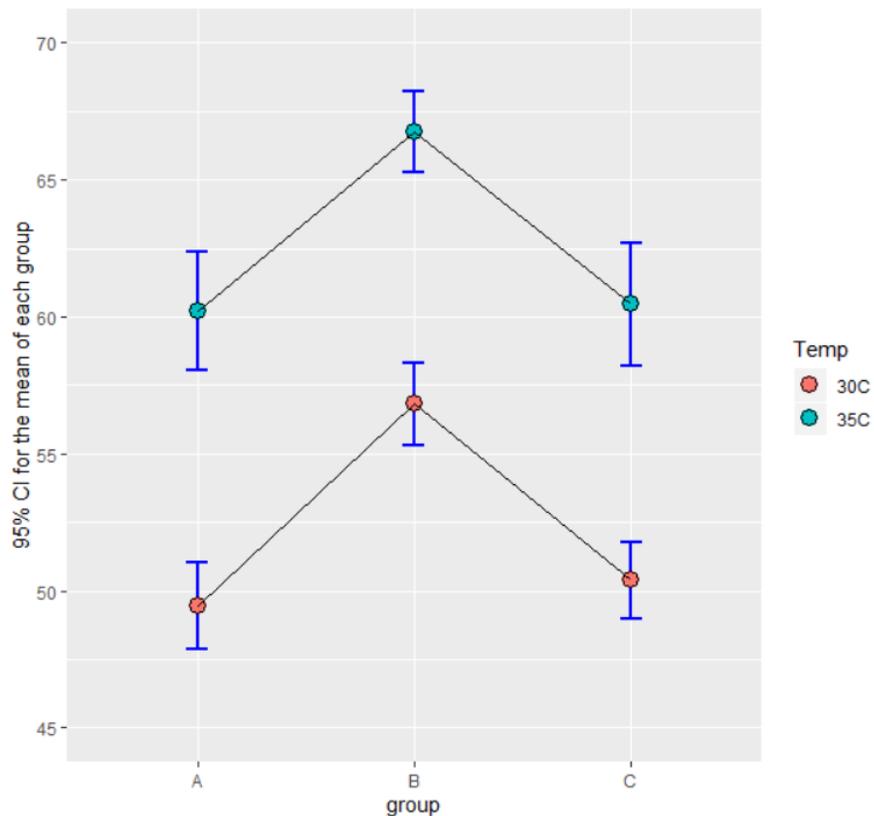
Diseño con dos factores

```
res <- lm(biomarker~group*Temp, data=dt)
anova(res)
```

Analysis of Variance Table

Response: biomarker

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
group	2	747.47	373.74	46.0673	1.96e-13	***
Temp	1	1979.13	1979.13	243.9505	< 2.2e-16	***
group:Temp	2	2.59	1.30	0.1597	0.8527	
Residuals	69	559.79	8.11			



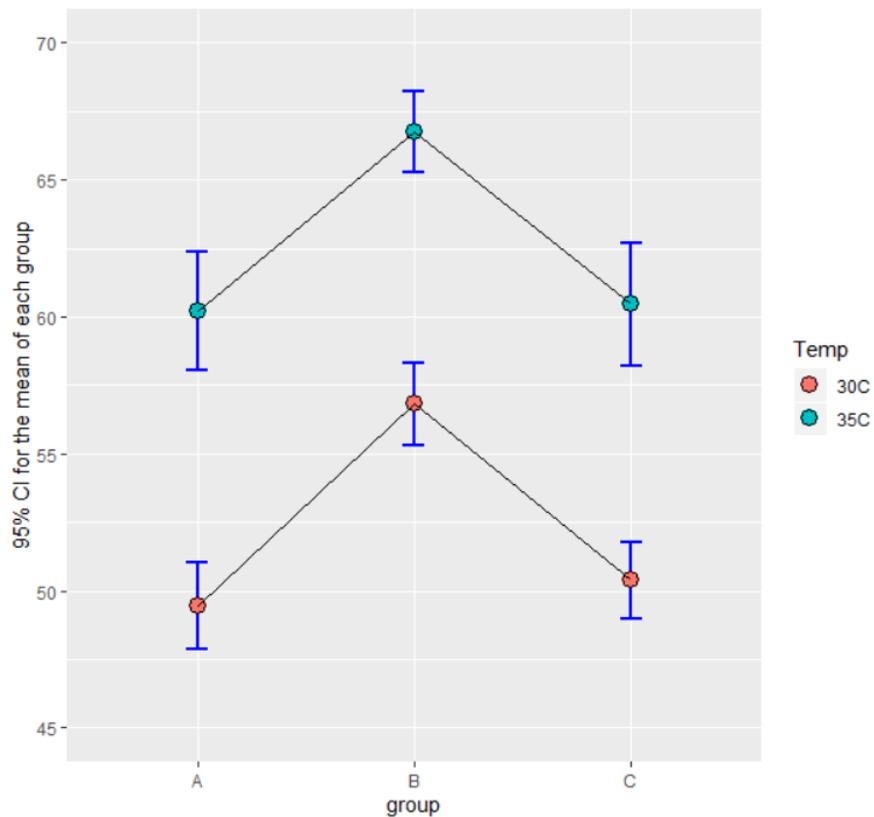
La interacción no es significativa. En la gráfica se interpreta que la diferencia de producción para cualquier substrato es mayor a 35°C y que esta diferencia es muy similar para cada substrato.

En este caso, deberíamos ajustar el modelo sin interacción. El resultado se muestra en la diapositiva siguiente.

Ejemplo

Diseño con dos factores

```
res <- lm(biomarker~group+Temp,data=dt)
anova(res)
```



Analysis of Variance Table

Response: biomarker

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
group	2	747.47	373.74	47.184	9.218e-14	***
Temp	1	1979.13	1979.13	249.865	< 2.2e-16	***
Residuals	71	562.38	7.92			

```
> confint(res)
```

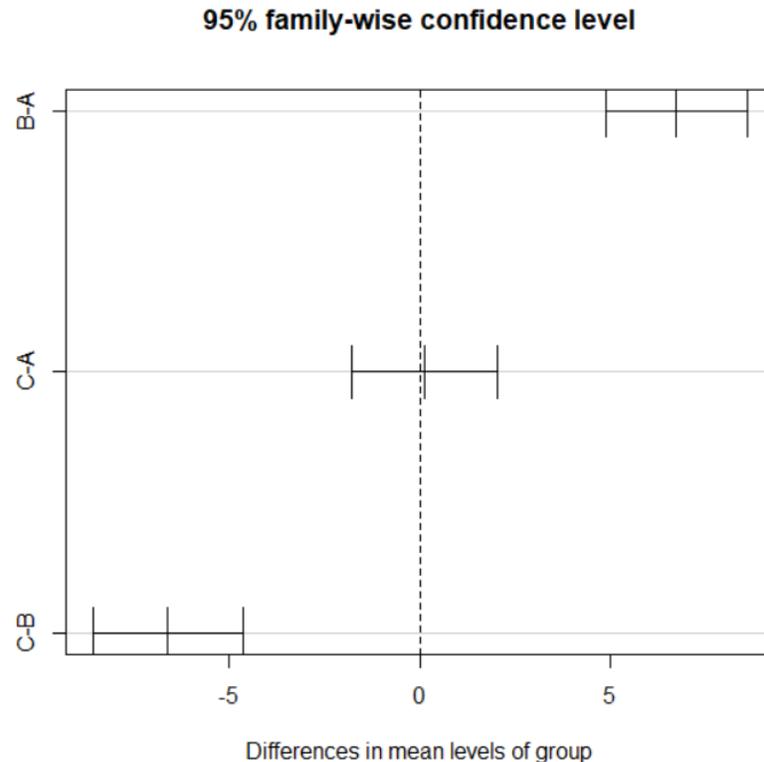
	2.5 %	97.5 %
(Intercept)	48.444249	50.930881
groupB	5.413283	8.501998
groupC	-1.005244	2.194518
Temp35C	8.991259	11.587053

El aumento medio de producción a 35C respecto a 30C para cualquier substrato está entre 8.99 i 11.59

Ejemplo

Diseño con dos factores

```
TukeyHSD(aov(res), 'group')  
TukeyHSD(aov(res), 'group') %>% plot()
```



Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = res)

\$group	diff	lwr	upr	p adj
B-A	6.7518571	4.898038	8.605677	0.0000000
C-A	0.1269481	-1.792486	2.046382	0.9862786
C-B	-6.6249091	-8.594364	-4.655454	0.0000000

El substrato B produce de media más que A y C. Estos dos últimos tienen una producción similar.

Los IC son amplios, por lo que convendría aumentar la muestra.

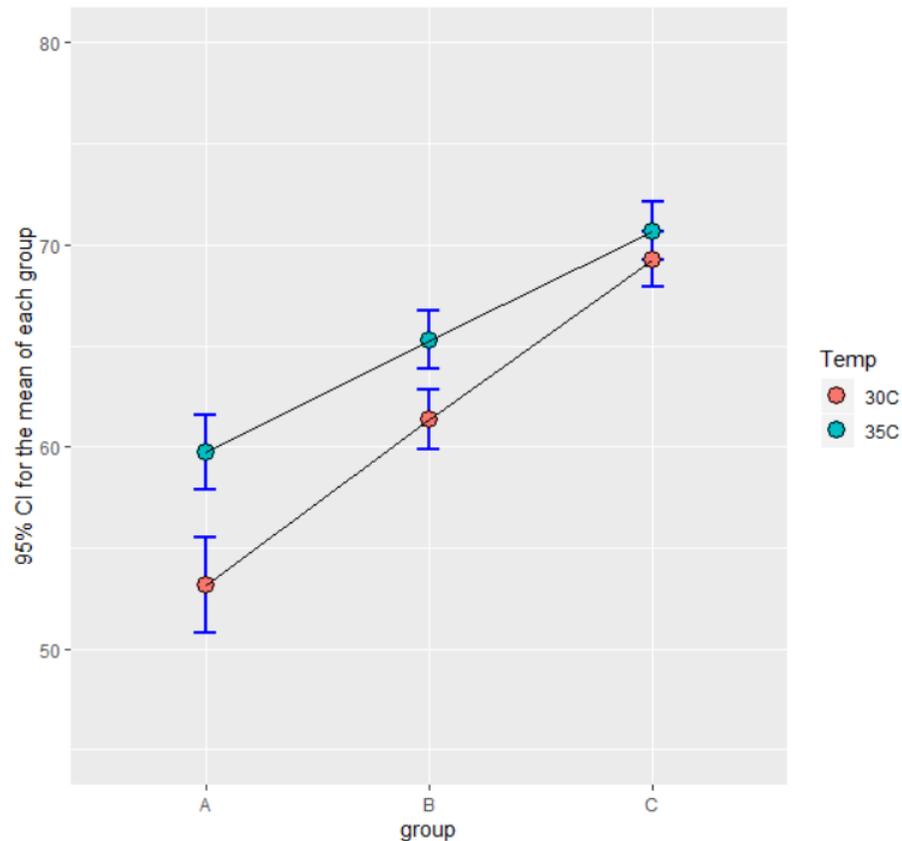
El concepto de interacción

Analysis of Variance Table

Response: biomarker

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
group	2	1777.39	888.69	134.583	< 2.2e-16	***
Temp	1	236.26	236.26	35.778	8.792e-08	***
group:Temp	2	70.01	35.00	5.301	0.007218	**
Residuals	69	455.63	6.60			

La interacción significa, en este caso, que el efecto de la temperatura no es el mismo para cada sustrato.



```
$`group:Temp`
```

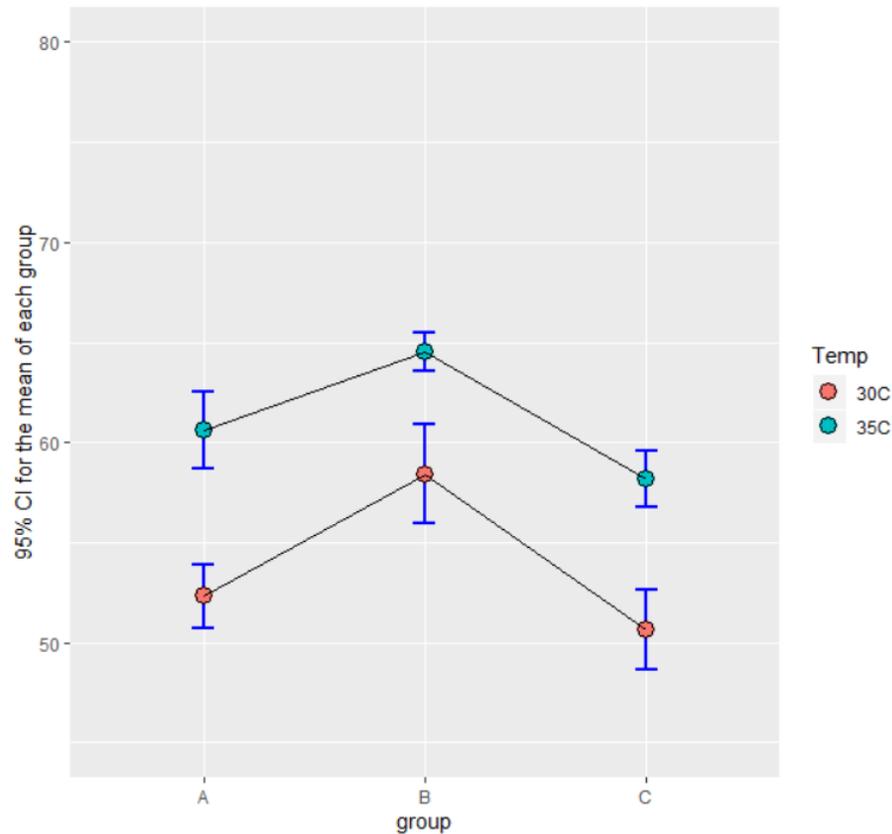
	diff	lwr	upr	p adj
B:30C-A:30C	8.221429	4.734511	11.708347	0.0000000
C:30C-A:30C	16.133613	12.750806	19.516421	0.0000000
A:35C-A:30C	6.577143	2.865035	10.289250	0.0000287
B:35C-A:30C	12.150893	8.737392	15.564394	0.0000000
C:35C-A:30C	17.520779	13.878812	21.162746	0.0000000
C:30C-B:30C	7.912185	5.193634	10.630736	0.0000000
A:35C-B:30C	-1.644286	-4.763080	1.474509	0.6365152
B:35C-B:30C	3.929464	1.172814	6.686115	0.0011501
C:35C-B:30C	9.299351	6.264378	12.334323	0.0000000
A:35C-C:30C	-9.556471	-12.558414	-6.554527	0.0000000
B:35C-C:30C	-3.982721	-6.606442	-1.359000	0.0004476
C:35C-C:30C	1.387166	-1.527598	4.301929	0.7298224
B:35C-A:35C	5.573750	2.537260	8.610240	0.0000139
C:35C-A:35C	10.943636	7.652405	14.234868	0.0000000
C:35C-B:35C	5.369886	2.419556	8.320217	0.0000166

El concepto de interacción

Analysis of Variance Table

Response: biomarker

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
group	2	629.96	314.98	38.5507	5.751e-12	***
Temp	1	1021.04	1021.04	124.9659	< 2.2e-16	***
group:Temp	2	15.00	7.50	0.9177	0.4043	
Residuals	69	563.77	8.17			



Analysis of Variance Table

Response: biomarker

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
group	2	629.96	314.98	38.64	4.428e-12	***
Temp	1	1021.04	1021.04	125.26	< 2.2e-16	***
Residuals	71	578.76	8.15			

\$group

	diff	lwr	upr	p adj
B-A	4.991883	3.044684	6.9390818	0.0000001
C-A	-2.205571	-4.086207	-0.3249363	0.0174961
C-B	-7.197455	-9.195398	-5.1995110	0.0000000

\$Temp

	diff	lwr	upr	p adj
35C-30C	7.394302	6.076648	8.711956	0

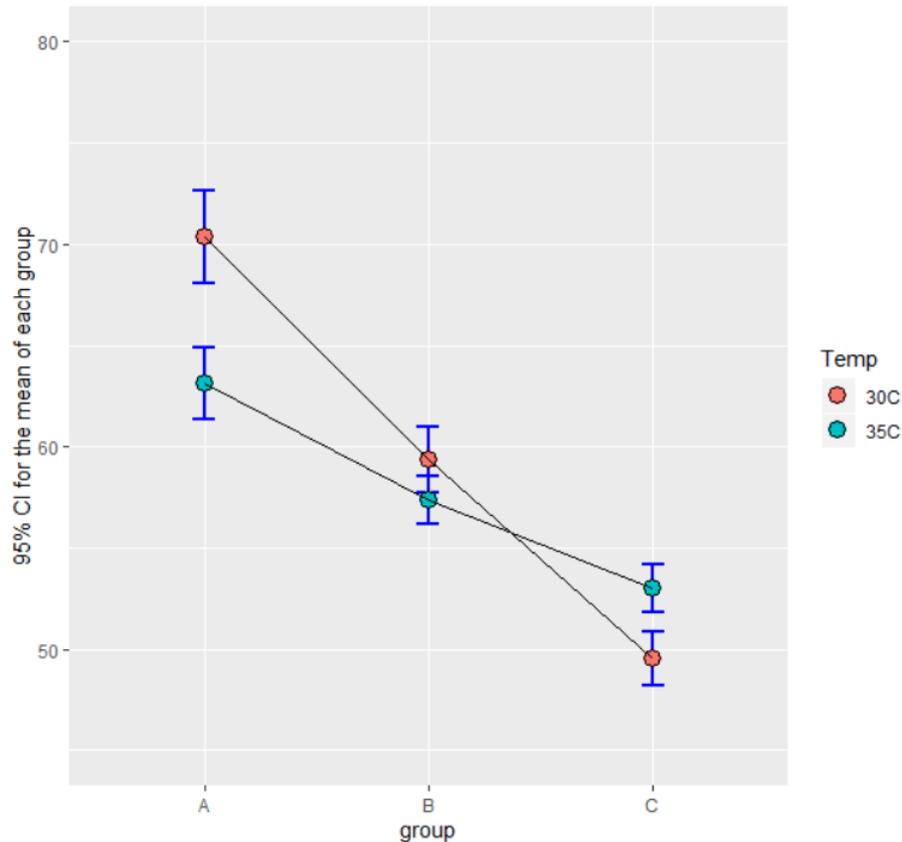
En este caso no hay interacción. El efecto de la temperatura (IC 6.1, 8.7) es el mismo en cada sustrato. Los sustratos A y C dan resultados similares para cada temperatura. El sustrato B da una producción media superior.

El concepto de interacción

Analysis of Variance Table

Response: biomarker

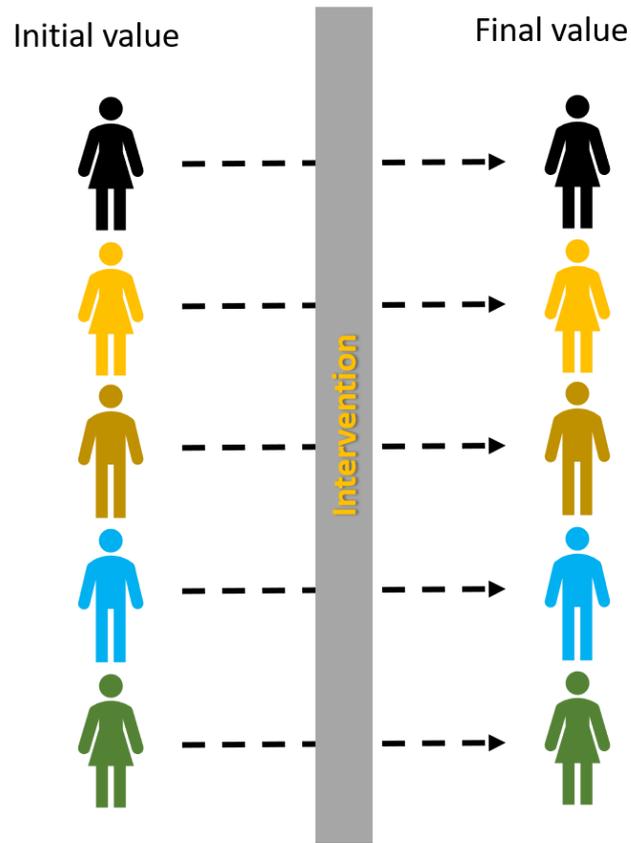
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
group	2	2553.98	1276.99	196.5678	< 2.2e-16	***
Temp	1	40.29	40.29	6.2015	0.01518	*
group:Temp	2	327.57	163.79	25.2117	6.035e-09	***
Residuals	69	448.25	6.50			



	diff	lwr	upr	p adj
B:30C-A:30C	-10.981250	-14.2164517	-7.7460483	0.0000000
C:30C-A:30C	-20.803571	-24.1149060	-17.4922369	0.0000000
A:35C-A:30C	-7.189286	-10.5006203	-3.8779512	0.0000003
B:35C-A:30C	-12.970455	-16.4421066	-9.4988025	0.0000000
C:35C-A:30C	-17.308333	-20.7185354	-13.8981313	0.0000000
C:30C-B:30C	-9.822321	-12.5565659	-7.0880769	0.0000000
A:35C-B:30C	3.791964	1.0577198	6.5262088	0.0016866
B:35C-B:30C	-1.989205	-4.9155546	0.9371455	0.3571246
C:35C-B:30C	-6.327083	-9.1802631	-3.4739036	0.0000002
A:35C-C:30C	13.614286	10.7903701	16.4382013	0.0000000
B:35C-C:30C	7.833117	4.8228128	10.8434210	0.0000000
C:35C-C:30C	3.495238	0.5560136	6.4344626	0.0106612
B:35C-A:35C	-5.781169	-8.7914730	-2.7708647	0.0000053
C:35C-A:35C	-10.119048	-13.0582721	-7.1798231	0.0000000
C:35C-B:35C	-4.337879	-7.4566081	-1.2191495	0.0016203

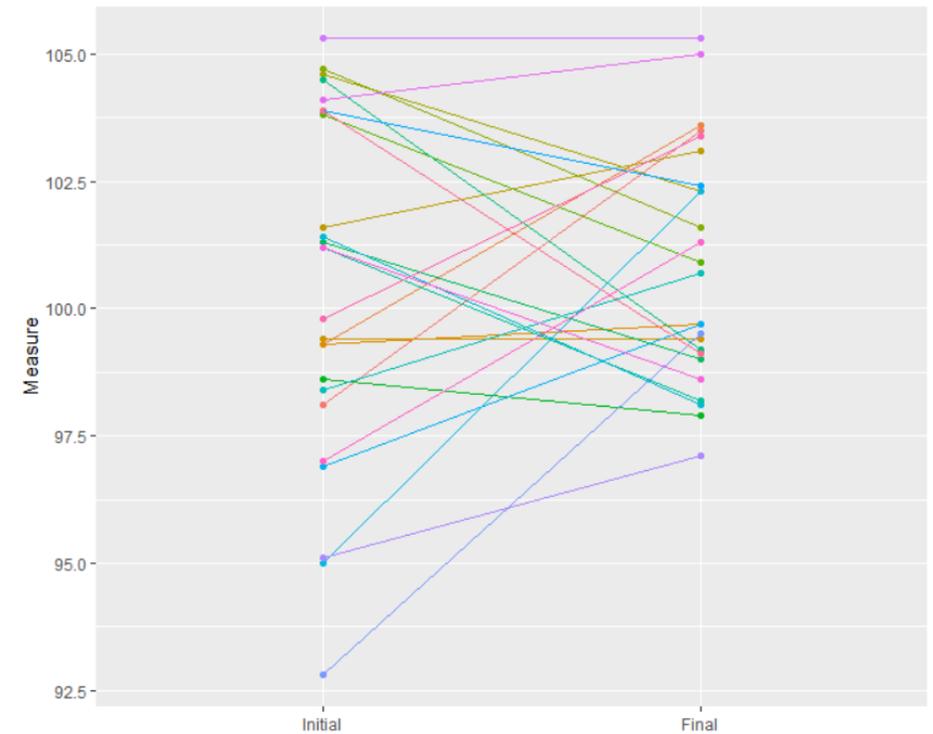
El la producción media en el sustrato A es superior a 30C (IC 3.9, 10.5) . Para B la producción es similar a ambas temperaturas (IC 0.9, -4.9). Para C, la producción es superior a 35C (IC 0.55, 6.4).

Diseño con datos apareados



init	final
98.1	103.5
99.3	103.6
99.3	99.7
99.4	99.4
101.6	103.1
104.6	102.3
104.7	101.6
103.8	100.9
98.6	97.9
101.3	99.0
104.5	99.2
101.2	98.2
98.4	100.7
101.4	98.1
95.0	102.3
96.9	99.7
103.9	102.4
92.8	99.5
95.1	97.1
105.3	105.3
104.1	105.0
101.2	98.6
97.0	101.3
99.8	103.4
103.9	99.1

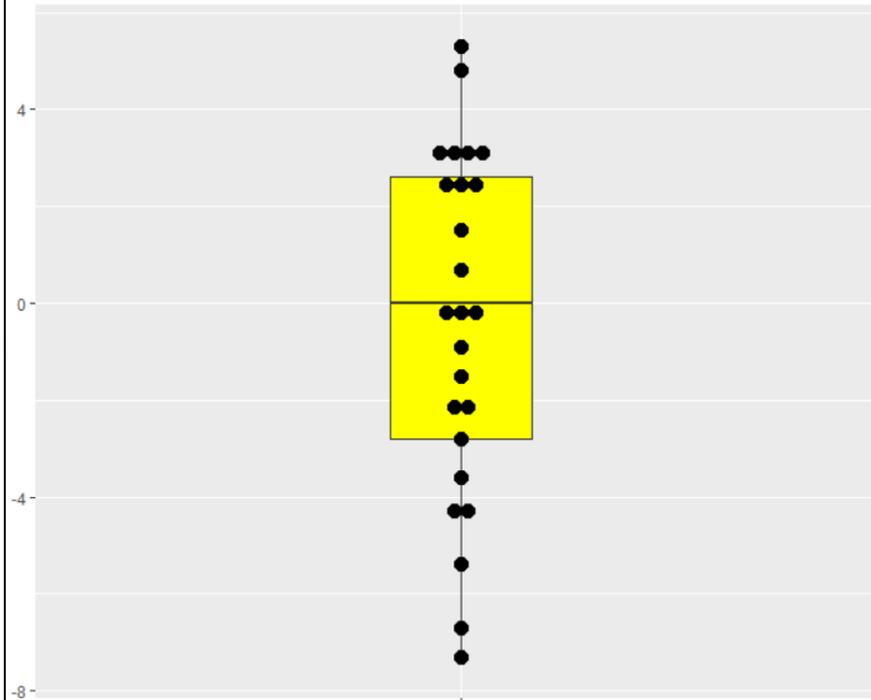
Cada individuo se mide antes y después del tratamiento. La evaluación del efecto se realiza estimado la media de la diferencia entre la situación inicial y final.



```
d <- d %>% mutate(dif=init-final)
```

init	final	dif
98.1	103.5	-5.4
99.3	103.6	-4.3
99.3	99.7	-0.4
99.4	99.4	0.0
101.6	103.1	-1.5
104.6	102.3	2.3
104.7	101.6	3.1
103.8	100.9	2.9
98.6	97.9	0.7
101.3	99.0	2.3
104.5	99.2	5.3
101.2	98.2	3.0
98.4	100.7	-2.3
101.4	98.1	3.3
95.0	102.3	-7.3
96.9	99.7	-2.8
103.9	102.4	1.5
92.8	99.5	-6.7
95.1	97.1	-2.0
105.3	105.3	0.0
104.1	105.0	-0.9
101.2	98.6	2.6
97.0	101.3	-4.3
99.8	103.4	-3.6
103.9	99.1	4.8

```
ggplot(d,aes(x='',y=dif))+  
  geom_boxplot(width=.2,fill='yellow')+  
  geom_dotplot(binaxis = 'y',  
              stackdir = 'center',  
              dotsize=.7,color='black')+  
  xlab('')
```



Cada individuo se mide antes y después del tratamiento. La evaluación del efecto se realiza estimado la media de la diferencia entre la situación inicial y final.

```
> t.test(d$dif)
```

One Sample t-test

```
data: d$dif
```

```
t = -0.546, df = 24, p-value = 0.5901
```

```
alternative hypothesis: true mean is not equal to 0  
95 percent confidence interval:
```

```
-1.854643  1.078643
```

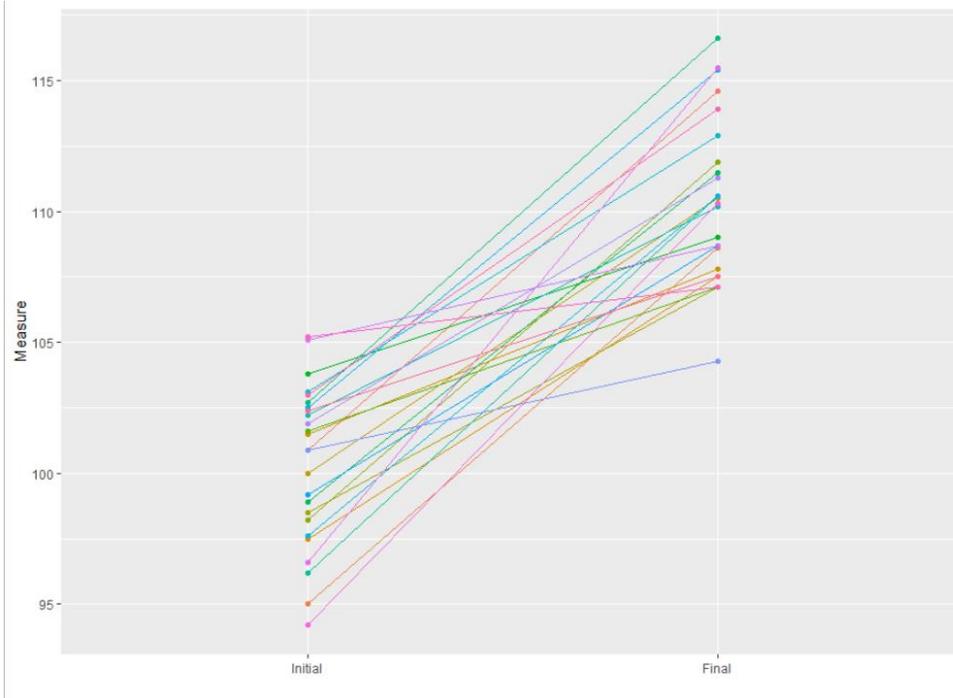
```
sample estimates:
```

```
mean of x  
-0.388
```

El IC de la diferencia entre los valores iniciales y finales (variable dif) es (-1.85, 1.08). Por lo tanto, no podemos descartar que la media de la diferencia sea 0.

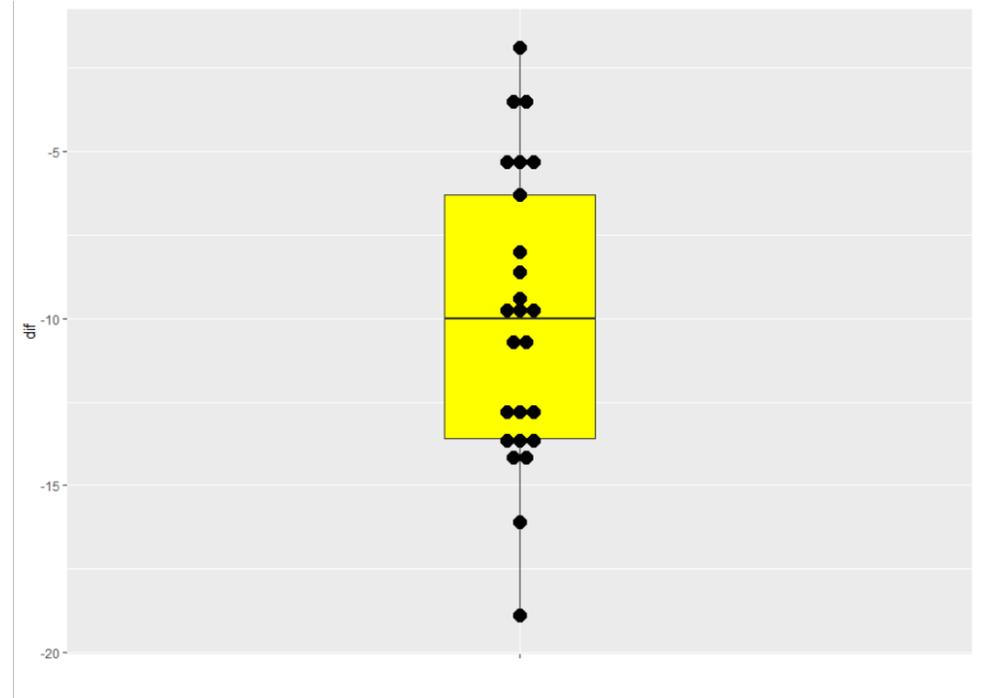
En este caso, el tratamiento no produciría un cambio apreciable en los valores medios.

Ejemplo



One Sample t-test

```
data: d$dif
t = -11.515, df = 24, p-value = 2.919e-11
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -11.815868  -8.224132
sample estimates:
mean of x
 -10.02
```



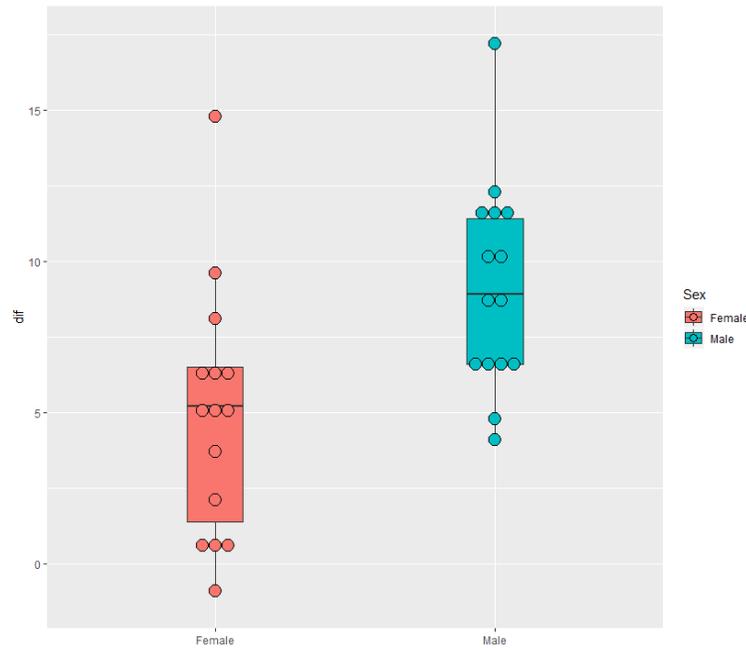
El IC de la media del incremento entre el final y el inicial es (8.22, 11.8).

Podemos concluir que se produce un aumento de entre 8.22 y 11.8.

¿Podemos concluir que el aumento es superior a 10? No, porque el IC incluye valores inferiores. Para llegar a esta conclusión el IC debería estar por encima de 10.

Ejemplo

init	final	Sex	dif
97.4	109.2	Male	11.8
104.2	113.1	Male	8.9
99.5	109.6	Male	10.1
101.9	113.3	Male	11.4
106.4	110.5	Male	4.1
99.9	112.2	Male	12.3
98.4	109.8	Male	11.4
104.0	110.9	Male	6.9
101.9	108.5	Male	6.6
99.9	110.1	Male	10.2
95.1	112.3	Male	17.2
97.0	105.5	Male	8.5
103.6	109.9	Male	6.3
101.4	108.0	Male	6.6
100.0	104.8	Male	4.8
100.7	105.9	Female	5.2
97.9	102.7	Female	4.8
103.8	104.3	Female	0.5
101.3	105.0	Female	3.7
97.5	107.1	Female	9.6
99.5	105.9	Female	6.4
100.9	100.0	Female	-0.9
94.4	102.5	Female	8.1
100.5	106.5	Female	6.0
104.3	105.0	Female	0.7
96.9	103.5	Female	6.6
102.7	103.2	Female	0.5
101.3	106.6	Female	5.3
94.0	108.8	Female	14.8
101.8	103.9	Female	2.1



Sin tener en cuenta Sex

One sample t-test

```
data: dif
t = 8.9355, df = 29, p-value = 7.951e-10
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 5.410644 8.622690
sample estimates:
mean of x
 7.016667
```

Sex	Mean	SD	N
Female	4.9	4.1	15
Male	9.1	3.4	15

```
> anova(res)
```

Analysis of Variance Table

Response: dif

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Sex	1	135.26	135.256	9.4395	0.004693 **
Residuals	28	401.21	14.329		

```
> TukeyHSD(aov(res))
```

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = res)

\$Sex	diff	lwr	upr	p adj
Male-Female	4.246667	1.41534	7.077993	0.0046928

La media de la diferencia entre el valor final y el inicial es superior en los hombres respecto a las mujeres (IC 1.4, 7.1).

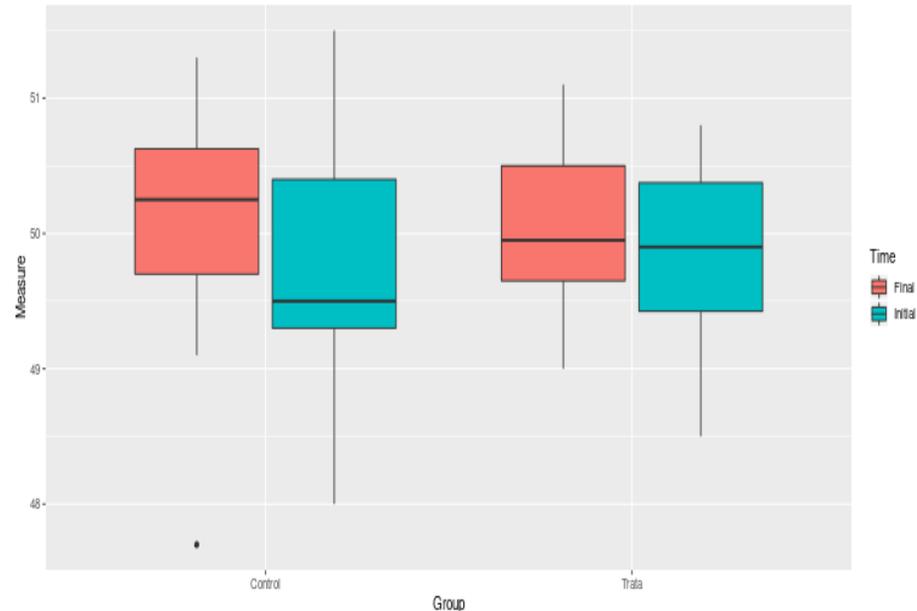
Efecto placebo y efecto del tratamiento

Diseñamos un experimento con un grupo control y un tratamiento y queremos estimar las cuestiones siguientes:

- Equivalencia entre el grupo control y el tratamiento al inicio.
- Estimación del efecto placebo.
- Estimación del efecto del tratamiento.

Group	Initial	Final	dif
Control	48.00	49.70	1.70
Control	49.30	49.10	-0.20
Control	49.30	50.20	0.90
Control	51.50	50.40	-1.10
Control	50.50	50.80	0.30
Control	50.10	49.70	-0.40
Control	49.40	47.70	-1.70
Control	50.50	50.70	0.20
Control	49.60	51.30	1.70
Control	49.20	50.30	1.10

Group	Initial	Final	dif
Trata	49.80	50.50	0.70
Trata	48.80	50.50	1.70
Trata	50.80	49.00	-1.80
Trata	49.60	49.10	-0.50
Trata	50.30	49.80	-0.50
Trata	50.00	50.10	0.10
Trata	50.60	51.10	0.50
Trata	49.70	49.90	0.20
Trata	48.50	50.00	1.50
Trata	48.90	49.90	1.00
Trata	50.60	49.20	-1.40
Trata	50.10	50.80	0.70



Equivalencia de condiciones iniciales

Para comparar si los grupos tienen valores medios similares, estimamos la diferencia de las medias poblacionales de los valores iniciales entre los grupos

El IC (-0.85, 0.72) contiene el valor cero e indica que los datos son compatibles con una igualdad de medias poblacionales entre los grupos.

```
t.test(Initial~Group,data=df)
```

```
Welch Two Sample t-test

data:  Initial by Group
t = -0.18367, df = 16.971, p-value = 0.8564
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.8533792  0.7167125
sample estimates:
mean in group Control    mean in group Trata
          49.74000              49.80833
```

Efecto placebo

Para estimar el efecto placebo, estimaremos la media de la diferencia (final-inicial) en el grupo control. Con ello, estamos estimando si se produce algún cambio significativo asociado al hecho de estar tratado con un placebo.

El IC (-0.56, 1.06) contiene el valor cero e indica que el grupo control no varía en el valor medio de la diferencia entre la situación final y la inicial.

Por lo tanto, concluimos que no hay un efecto placebo apreciable.

```
df <- df %>% filter(Group=='Control')  
  
with(df,t.test(dif))
```

```
One Sample t-test  
  
data: dif  
t = 0.69621, df = 9, p-value = 0.5039  
alternative hypothesis: true mean is not equal to 0  
95 percent confidence interval:  
-0.5623143 1.0623143  
sample estimates:  
mean of x  
0.25
```

Evolución del grupo de tratamiento

Para estimar la evolución del grupo de tratamiento, estimaremos la media de la diferencia (final-inicial) en el grupo de tratamiento. Con ello, estamos estimando si se produce algún cambio significativo asociado al hecho de estar tratado.

El IC (-0.50, 0.87) contiene el valor cero e indica que el grupo de tratamiento no varía en el valor medio de la diferencia entre la situación final y la inicial.

Por lo tanto, concluimos que no hay un efecto del tratamiento.

```
df <- df %>% filter(Group=='Trata')
```

```
with(df, t.test(dif))
```

One Sample t-test

```
data: dif
t = 0.59067, df = 11, p-value = 0.5667
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.4998178  0.8664845
sample estimates:
mean of x
0.1833333
```

Estimación del efecto del tratamiento

Para estimar el efecto del tratamiento, estimaremos la diferencia entre la evolución del tratamiento y del control. Con ello, estamos estimando si la evolución del control es superior a la observada en el control (efecto placebo).

El IC (-1.06, 0.93) contiene el valor cero e indica que no podemos concluir que la evolución del grupo de tratamiento sea superior a la del grupo control.

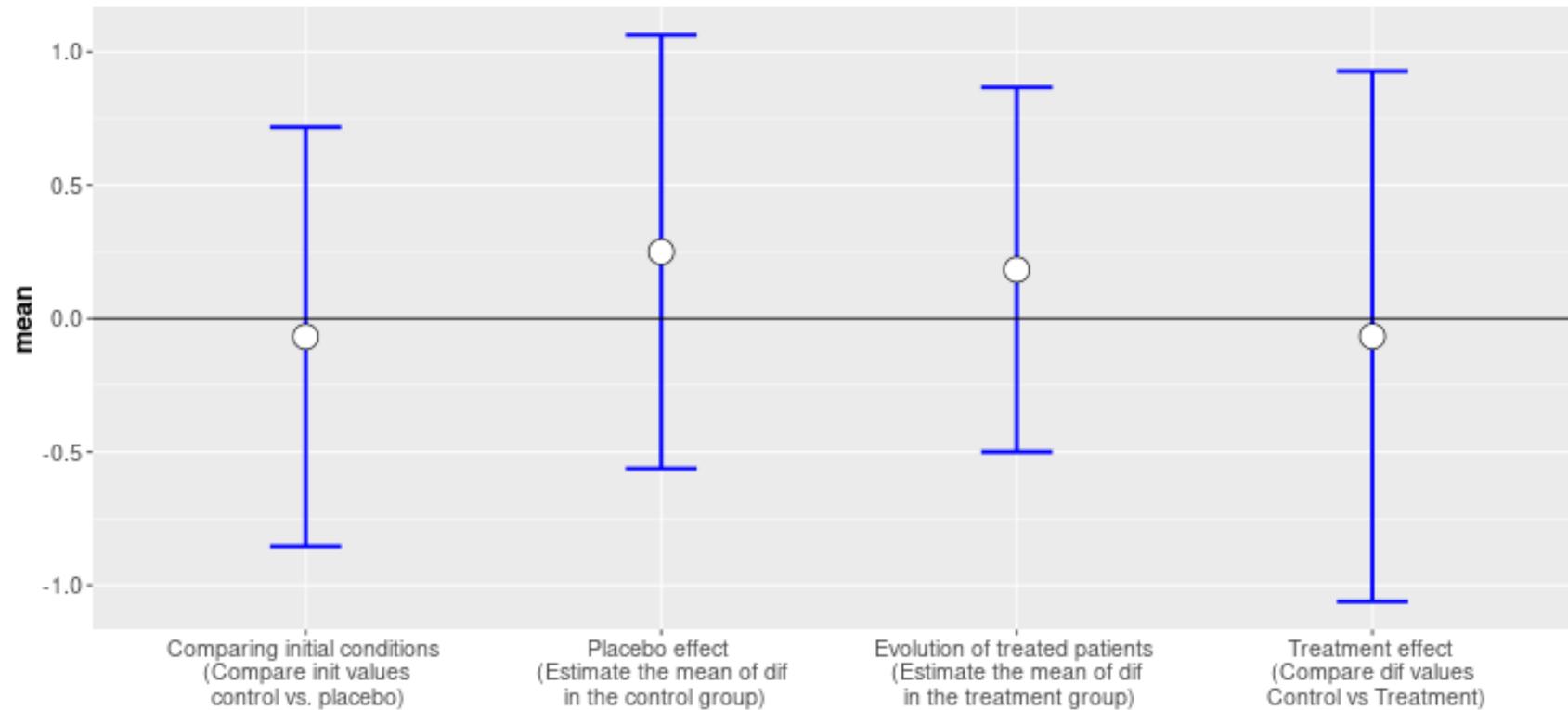
Por lo tanto, concluimos que no hay un efecto del tratamiento.

```
t.test(dif~Group,data=df)
```

```
Welch Two Sample t-test

data:  Tdif and Cdif
t = -0.14046, df = 18.859, p-value = 0.8898
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.0606018  0.9272684
sample estimates:
mean of x mean of y
0.1833333 0.2500000
```

Comparación de los resultados



Estimación del efecto

Ejemplo donde existe efecto placebo

- (a) Comparación condiciones iniciales
- (b) Efecto placebo
- (c) Evolución tratamiento
- (d) Efecto del tratamiento (descontando efecto placebo)

```
t.test(Initial~Group,data=df)
```

```
Welch Two Sample t-test (a)

data: Initial by Group
t = -1.3319, df = 19.198, p-value = 0.1985
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.3794005  0.3060672
sample estimates:
mean in group Control  mean in group Trata
          49.53000          50.06667
```

```
df <- df %>% filter(Group=='Trata')
```

```
with(df,t.test(dif))
```

```
One Sample t-test (c)

data: dif
t = 24.219, df = 11, p-value = 6.797e-11
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 5.416855 6.499811
sample estimates:
mean of x
5.958333
```

```
df <- df %>% filter(Group=='Control')
```

```
with(df,t.test(dif))
```

```
One Sample t-test (b)

data: dif
t = 10.973, df = 9, p-value = 1.643e-06
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 3.230966 4.909034
sample estimates:
mean of x
4.07
```

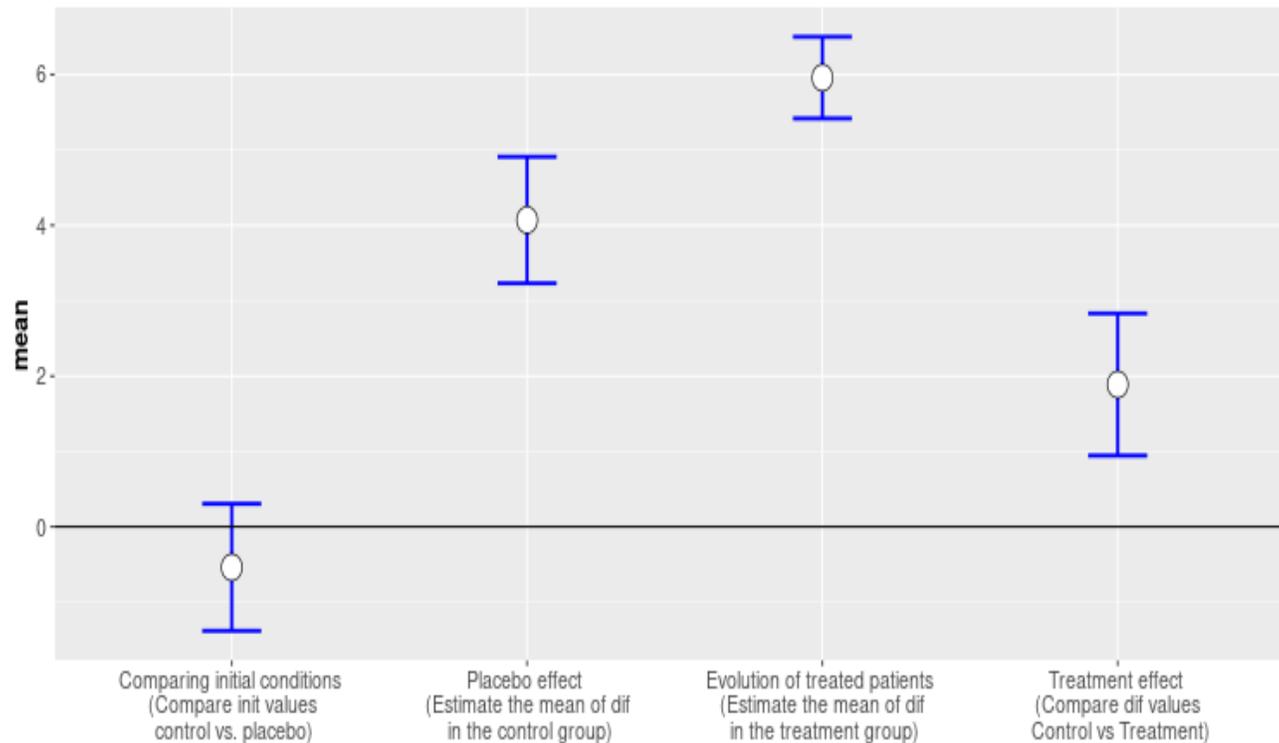
```
t.test(dif~Group,data=df)
```

```
Welch Two Sample t-test (d)

data: Tdif and Cdif
t = 4.2427, df = 16.11, p-value = 0.0006119
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.9453425 2.8313242
sample estimates:
mean of x mean of y
5.958333 4.070000
```

Estimación del efecto

Ejemplo donde existe efecto placebo



Los valores iniciales tiene medias similares en ambos grupos.

La evolución del grupo control muestra un efecto placebo importante.

La evolución del grupo de tratamiento muestra un efecto importante entre 5.4 y 6.5

Cuando se comparan los valores de evolución (dif) entre ambos grupos, se estima el efecto del tratamiento descontando el efecto del placebo. En este caso, el efecto el tratamiento estaría entre 0.95 y 2.83.

Estimación del efecto

Ejemplo donde existe efecto placebo

- (a) Comparación condiciones iniciales
- (b) Efecto placebo
- (c) Evolución tratamiento
- (d) Efecto del tratamiento (descontando efecto placebo)

```
t.test(Initial~Group,data=df)
```

```
Welch Two Sample t-test (a)

data: Initial by Group
t = -1.2122, df = 19.938, p-value = 0.2396
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.2245021  0.3245021
sample estimates:
mean in group Control    mean in group Trata
                50.15                50.60
```

```
df <- df %>% filter(Group=='Trata')
```

```
with(df,t.test(dif))
```

```
One Sample t-test (c)

data: dif
t = 7.394, df = 11, p-value = 1.37e-05
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 2.586906 4.779761
sample estimates:
mean of x
3.683333
```

```
df <- df %>% filter(Group=='Control')
```

```
with(df,t.test(dif))
```

```
One Sample t-test (b)

data: dif
t = 8.129, df = 9, p-value = 1.948e-05
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 2.634266 4.665734
sample estimates:
mean of x
3.65
```

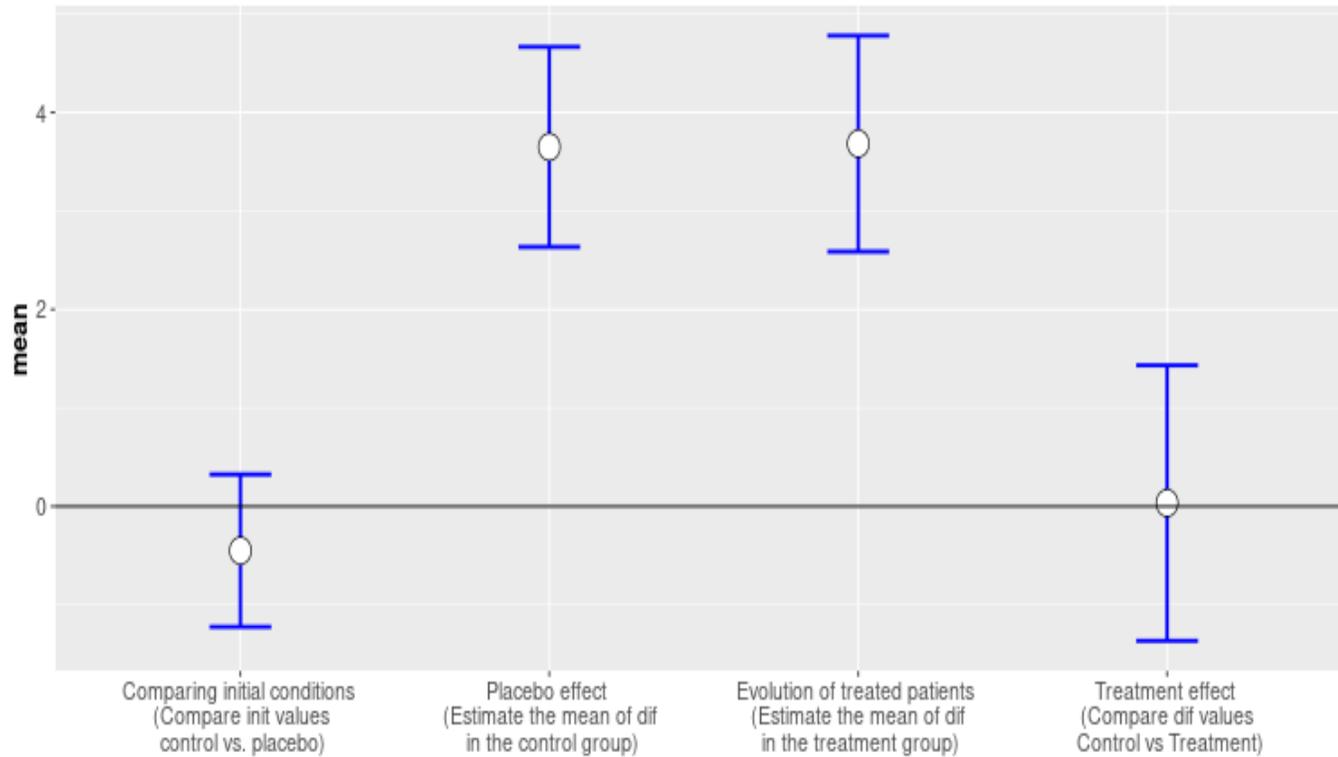
```
t.test(dif~Group,data=df)
```

```
Welch Two Sample t-test (d)

data: Tdif and Cdif
t = 0.049703, df = 20, p-value = 0.9609
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.365613  1.432280
sample estimates:
mean of x mean of y
3.683333  3.650000
```

Estimación del efecto

Ejemplo donde existe efecto placebo



Los valores iniciales tiene medias similares en ambos grupos.

La evolución del grupo control muestra un efecto placebo importante.

La evolución del grupo de tratamiento muestra un efecto importante (pero similar al placebo).

Cuando se comparan los valores de evolución (*dif*) entre ambos grupos, se estima el efecto del tratamiento descontando el efecto del placebo. En este caso, el efecto el tratamiento es inexistente (IC -1.4,1.4).

Conclusiones

El análisis de la varianza (ANOVA) es una herramienta básica para la evaluación de resultados en experimentos donde la variable respuesta es cuantitativa.

La tabla de ANOVA evalúa la significación de los factores y de su efecto en las medias de la variable respuesta.

El procedimiento **TukeyHSD** estima las diferencias de medias entre distintos niveles de un factor.

Cuando hay más de un factor, es importante verificar si existe interacción, es decir si las diferencias debidas a un factor son iguales o no para los distintos niveles de otro factor.

Es importante utilizar un diseño apropiado para estimar el efecto de un tratamiento y considerar el posible efecto placebo.

¿Qué debemos entender?

Diseños experimentales básicos

- Control vs. Tratamiento (t.test)
- Control vs. Varios tratamientos (lm, anova, TukeyHSD)
- Control vs. Varios tratamientos y un factor adicional (lm, anova, TukeyHSD)
- Interpretación de la tabla de ANOVA
- Interpretación de los intervalos proporcionados por TukeyHSD
- Funciones de R para este tipo de análisis

Concepto de efecto placebo

- Diseño básico para estimar el efecto placebo y el efecto del tratamiento
- Procedimientos para estimar los distintos efectos
- Interpretación
- Funciones de R para este tipo de análisis