

Tutorial para determinação de concentrações letais e Rácio de resistência em R com suporte do Rstudio

Bruno Gomes

Laboratório de Bioquímica e Fisiologia de Insetos (LABFISI)
Instituto Oswaldo Cruz (IOC-Fiocruz)
Rio de Janeiro, Brasil

Este tutorial está acompanhado por três ficheiros de apoio e tem como funcionalidade determinar parâmetros importantes para avaliar a ação de substâncias com propriedades inseticidas em insetos vetores. Estes métodos também podem ser aplicados em outros cenários de toxicologia.

Os ficheiros anexados a este tutorial:

- R script chamado **Tutorial_LC.R**
- R script chamado **Tutorial_RR.R**
- Ficheiro de texto chamado **input_tutorial.txt**

1. Aspectos básicos da utilização do R no Rstudio.

Caso já tenha conhecimento básico da utilização do R no Rstudio pode saltar a etapa 1 deste tutorial. As especificações sobre a análise para determinar concentrações letais é apenas iniciada na etapa 2.

1.1 Instalar R

O R atualiza a sua plataforma frequentemente e recomendamos instalar a versão mais atualizada. Alguns pacotes podem deixar de funcionar com novas versões do R. Este protocolo foi testado com sucesso na versão R 4.2.1 para Windows/Mac.

Link para download: <https://cran.r-project.org/bin/windows/base/>

1.2 Instalar Rstudio

Este é um software comercial (com versão básica gratuita – “Open Source”) que apoia uma melhor visualização do código e conteúdo quando trabalhamos em R.

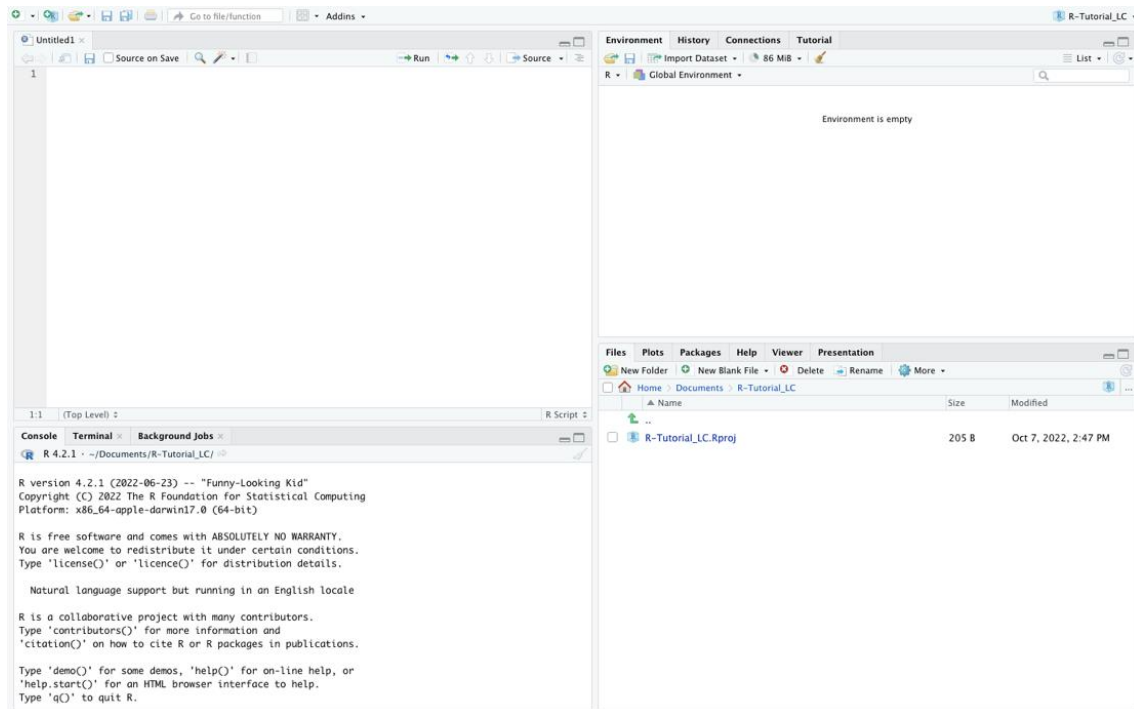
Link para download: <https://www.rstudio.com/products/rstudio/>

1.3 Rstudio e preparar ambiente de trabalho com diretório específico (pasta de trabalho)

Ao abrir o Rstudio é possível observar 4 quadrantes que normalmente estão organizados em *Scripts* no quadrante superior esquerdo, *Console* (e outras opções de leitura de código como o

terminal) no quadrante inferior esquerdo, elementos na memória virtual (*Environment*) e outras opções como o histórico da corrida no quadrante superior direito, enquanto no quadrante inferior direito temos a visualização de gráficos e opções de apoio do R.

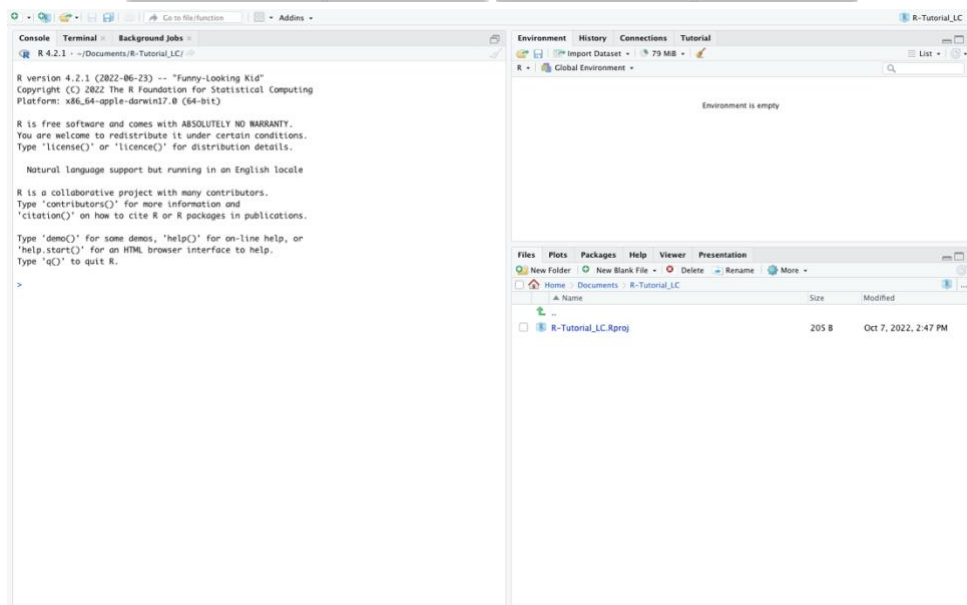
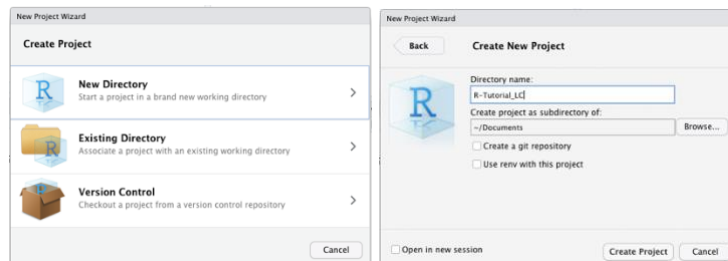
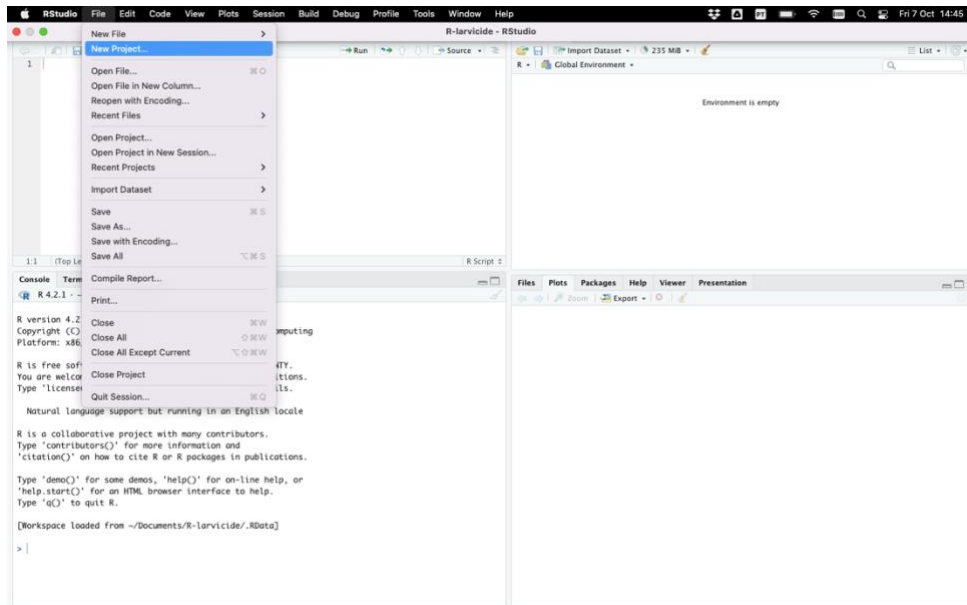
Esta organização e outras características de visualização do Rstudio podem ser ajustadas ao gosto do utilizador. Este tutorial não foca nestas alterações, mas caso esteja interessado pode explorar as opções em “Tools > Global Options”.



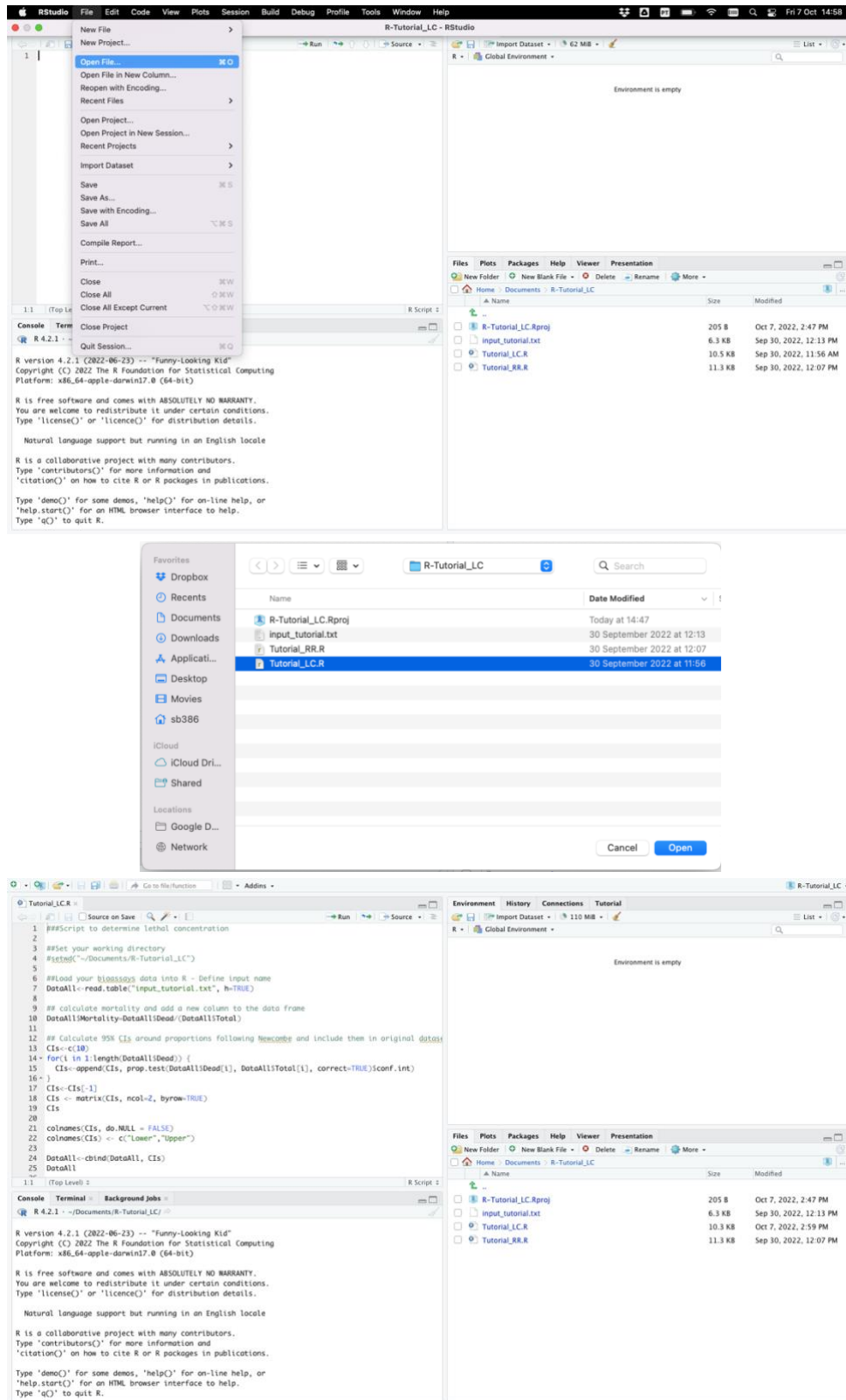
2.

2.1 Preparação do ambiente R no Rstudio para correr a análise

Para iniciar a atividade de análise no Rstudio deve-se criar um projeto (File > New Project) com o nome Tutorial_LC, onde o utilizador tem de definir um diretório específico que pode designar num diretório fácil de localizar (e.g. `~/Documents/R-Tutorial_LC`). O utilizador deve colocar os *Scripts* com o nome *Tutorial_LC.R* e *Tutorial_RR.R* e o ficheiro *input* com o nome *input_tutorial.txt* no diretório de trabalho definido. Todos os ficheiros criados pelo *Script* também serão armazenados neste diretório.



Este tutorial é dividido em duas abordagens que podem ser aplicadas em cenários diferentes. A abordagem 1 utiliza o *Script Tutorial_LC.R* e permite a análise de 3 populações sem definir uma linhagem de referência, onde não é possível calcular o rácio de resistência (Ponto 3). Enquanto, a abordagem 2 usa o *Script Tutorial_RR.R* para determinar o rácio de resistência de 2 populações em relação a um linhagem de referência (Ponto 4).



2.2 Dicas básicas para observar o Script

(1) A informação de apoio aparece em cor verde após o símbolo # (e.g. **#Set your working directory, #Definir concentrações**), e pode ser em qualquer idioma.

(2) As linhas de comando a serem lidas pela “Console” está normalmente na cor preta, mas existem algumas informações em destaque em outras cores, como os rótulos necessários para encontrar dados ou definir o nome do ficheiro a ser criado de cor verde entre parêntesis (e.g. “input”, “results.txt”), ou os parâmetros para correr a linha de comando na cor azul (e.g. **TRUE**).

Exemplo:

#Ler a informação do ficheiro input

DataAll<-read.table("input_tutorial.txt", h=TRUE**)**

Ou

DataAll<-read.table("input_tutorial.txt", h=TRUE**) #Ler a informação do ficheiro input**

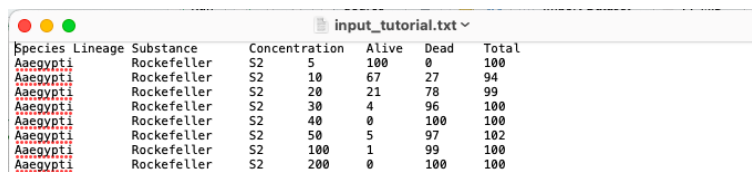
(3) O atalho “Ctrl + Enter” ordena a leitura da linha de comando onde o cursor se localiza.

3. Abordagem 1: cálculo de concentrações letais sem linhagem de referência

3.1 Leitura do ficheiro input

O R permite a leitura de diversos tipos de ficheiros de texto. Aqui estamos a fazer uma leitura de informação tabelada em um ficheiro básico de texto “.txt”. O parâmetro “**h=TRUE**” indica que o documento tem cabeçalho que é essencial para localização de informação.

O ficheiro *input* neste tutorial tem uma tabela com 7 colunas com os seguintes cabeçalhos: **Species, Lineage, Substance, Concentration, Alive, Dead, Total**. Em futuras utilizações, nós recomendamos a manutenção dos cabeçalhos para evitar alterações desnecessárias no *Script* original, mas caso alterem alguma coluna a informação do cabeçalho não deve conter acentos ou símbolos que não sejam do alfabeto padrão ou números.



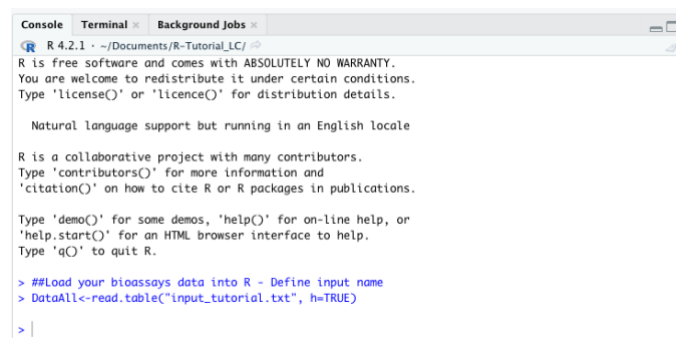
Species	Lineage	Substance	Concentration	Alive	Dead	Total
Aaegypti	Rockefeller	S2	5	100	0	100
Aaegypti	Rockefeller	S2	10	67	27	94
Aaegypti	Rockefeller	S2	20	21	78	99
Aaegypti	Rockefeller	S2	30	4	96	100
Aaegypti	Rockefeller	S2	40	0	100	100
Aaegypti	Rockefeller	S2	50	5	97	102
Aaegypti	Rockefeller	S2	100	1	99	100
Aaegypti	Rockefeller	S2	200	0	100	100

Este ficheiro deve conter informação de todos os ensaios realizados no experimento independentemente do número de substâncias ou espécies/linhagens envolvidas, visto que o *Script* está preparado para analisar múltiplos grupos em paralelo e pode ser ajustado conforme o número total de grupos. O nome do ficheiro pode ser ajustado conforme a necessidade do utilizador.

Linha de código:

```
DataAll<-read.table("input_tutorial.txt", h=TRUE)
```

Após a leitura deste comando os dados presentes no ficheiro serão armazenados na memória do projeto com o nome “DataAll” (esta terminologia é definida pelo utilizador, mas a sua alteração vai requerer alterar de forma consistente quando esta determinação aparecer no *Script*). A informação na memória do projeto pode ser observada no quadrante superior direito em *Environment*, onde deve aparecer que “DataAll” tem 192 observações (representando o número de linhas do ficheiro além do cabeçalho) com 7 variáveis obtidas das sete colunas definidas.



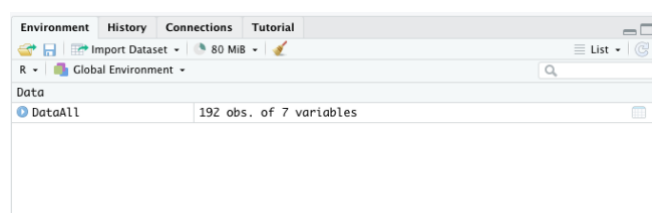
```
R 4.2.1 - ~/Documents/R-Tutorial_IC/
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

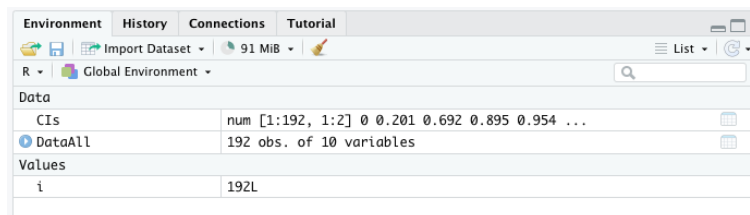
> ##Load your bioassays data into R - Define input name
> DataAll<-read.table("input_tutorial.txt", h=TRUE)
> |
```



3.2 Determinar mortalidade e intervalos de confiança a 95%

Nesta parte do Script, os dados de mortos e número totais de indivíduos são usados para determinar a mortalidade e realiza um loop para o cálculo dos intervalos de confiança. No final do script a informação calculada é incorporada na informação de “DataAll” criando três variáveis adicionais (**Mortality, Lower, Upper**).

```
9  ## calculate mortality and add a new column to the data frame
10 DataAll$Mortality=DataAll$Dead/(DataAll$Total)
11
12 ## Calculate 95% CIs around proportions following Newcombe and include them in original dataset
13 CIs<-c(10)
14 for(i in 1:length(DataAll$Dead)) {
15   CIs<-append(CIs, prop.test(DataAll$Dead[i], DataAll$Total[i], correct=TRUE)$conf.int)
16 }
17 CIs<-CIs[-1]
18 CIs <- matrix(CIs, ncol=2, byrow=TRUE)
19 CIs
20
21 colnames(CIs, do.NULL = FALSE)
22 colnames(CIs) <- c("Lower","Upper")
23
24 DataAll<-cbind(DataAll, CIs)
25 DataAll
26
```



Environment	
History	Connections
Tutorial	
R - Global Environment	
Data	
CIs	num [1:192, 1:2] 0 0.201 0.692 0.895 0.954 ...
DataAll	192 obs. of 10 variables
Values	
i	192L

3.3 Gerar bases de dados específicas dos subgrupos em análise.

A partir deste ponto, o tutorial vai calcular as concentrações letais para a substância S1 de três populações selvagens (duas populações do mosquito *Aedes aegypti*, e uma população do mosquito *Aedes albopictus*).

```
G1=DataAll[DataAll$Species=="Aegypti" &
           DataAll$Lineage=="Pop01" &
           DataAll$Substance=="S1" ,]
```

Esta linha de comando pode ser ajustada para selecionar diferentes populações baseada na informação de espécie, linhagem e/ou substância, e neste tutorial foi utilizado para fazer três seleções diferentes designadas por G1, G2 e G3.

```
27 ## create multiple data sets
28 G1=DataAll[DataAll$Species=="Aegypti" &
29           DataAll$Lineage=="Pop01" &
30           DataAll$Substance=="S1" ,]
31 G2=DataAll[DataAll$Species=="Aegypti" &
32           DataAll$Lineage=="Pop02" &
33           DataAll$Substance=="S1" ,]
34 G3=DataAll[DataAll$Species=="Aalbopictus" &
35           DataAll$Lineage=="Pop01" &
36           DataAll$Substance=="S1" ,]
37
```

Se verificar o *Environment* podemos verificar que os novos grupos definidos têm 24 observações. Caso exista alguma seleção que tenha “0 observações”, isso implica que houve um erro a selecionar os dados e o utilizador deve verificar se há algum erro no código (e.g. nome de algum dos items).

Environment	History	Connections	Tutorial
R - Global Environment			
Data			
CIs	num [1:192, 1:2] 0 0.201 0.692 0.895 0.954 ...		
DataAll	192 obs. of 10 variables		
G1	24 obs. of 10 variables		
G2	24 obs. of 10 variables		
G3	24 obs. of 10 variables		
Values			
i	192L		

3.4 Gerar modelo linear baseado em dados binomiais baseado na informação individual de mortalidade

Nesta etapa, o script cria vários modelos binomial logit usando os dados individuais de mortalidade de cada grupo definido anteriormente. Modelos binomiais probit também podem ser aplicados nesta fase.

```

48 # fit logistic regression model for multiple conditions
49 G1.glm=glm(data=G1,
50            cbind(Dead,Alive)~log(Concentration),
51            family="binomial"(link="logit"))

```

Environment	History	Connections	Tutorial
R - Global Environment			
Data			
CIs	num [1:192, 1:2] 0 0.201 0.692 0.895 0.954 ...		
DataAll	192 obs. of 10 variables		
G1	24 obs. of 10 variables		
G1.glm	List of 30		
G2	24 obs. of 10 variables		
G2.glm	List of 30		
G3	24 obs. of 10 variables		
G3.glm	List of 30		
Values			
i	192L		

3.5 Cálculo de concentrações letais

O função “dose.p” do pacote “MASS” vai ser usado para determinar a dose (neste caso equivalente a concentração letal) com mortalidade a 50%, 95% e 95%. As linhas seguintes calculam os intervalos de confiança a 95% e o erro padrão associado aos valores de concentração letal. Este script cria uma matriz com todos os valores calculados.

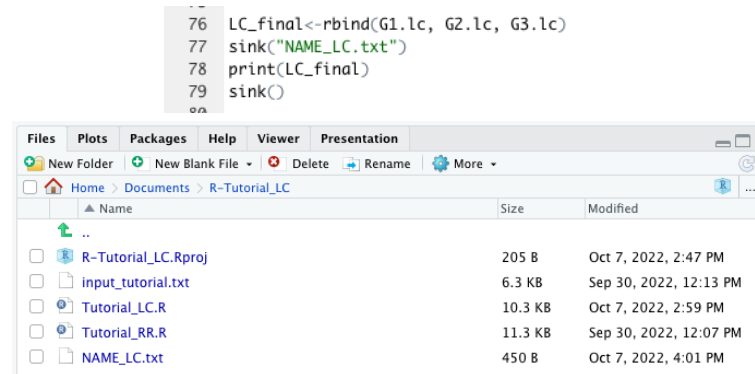
```

49 ##Calculate LCs for 3 populations - Install MASS
50 library(MASS)
51 G1.ld <- dose.p(G1.glm, p=c(0.50, 0.90, 0.95)) # from MASS
52 G1.ci <- G1.ld + attr(G1.ld, "SE") %*% matrix(qnorm(1 - 0.05/2)*c(-1,1), nrow=1)
53 G1.lc <- exp(cbind(G1.ld, attr(G1.ld, "SE"), G1.ci[,1], G1.ci[,2]))
54 dimnames(G1.lc)[[2]] <- c("LC", "SE", "Lower", "Upper")
55 rownames(G1.lc, do.NULL = FALSE)
56 rownames(G1.lc)<-c("G1-50", "G1-90", "G1-95")
57 G1.lc

```

Environment	History	Connections	Tutorial
R - Global Environment			
Data			
CIs	num [1:192, 1:2] 0 0.201 0.692 0.895 0.954 ...		
DataAll	192 obs. of 10 variables		
G1	24 obs. of 10 variables		
G1.ci	num [1:3, 1:2] 1.59 2.7 3.06 1.72 2.9 ...		
G1.glm	List of 30		
G1.lc	num [1:3, 1:4] 5.23 16.43 24.26 1.03 1.05 ...		
G2	24 obs. of 10 variables		
G2.glm	List of 30		
G3	24 obs. of 10 variables		
G3.glm	List of 30		
Values			
G1.ld	'glm.dose' Named num [1:3] 1.65 2.8 3.19		
i	192L		

Esta parte do script compila os dados calculados dos três grupos e exporta estes dados em um documento de texto com o nome "NAME_LC.txt" criando um ficheiro no diretório de trabalho. O nome do ficheiro deve ser ajustado sempre que corram o script com dados próprios de forma a evitar confusão entre corridas.



3.6 Preparação de gráficos

Informações de cinco pacotes vão ser necessários para efetuar todos os cálculos necessários. Nesta etapa são determinados base de dados de concentração e mortalidade baseado nos modelos de forma a termos dados para desenhar uma representação gráfica. O valor "0" e "200" são os valores mínimos e máximos do intervalo enquanto o "0.1" representa o intervalo entre os pontos criados. Estes valores devem ser ajustados caso exista grande variação de concentração entre substâncias/condições a serem analisadas.

```

82 ##Prepare graphic visualization - install multiple packages
83 library(readr)
84 library(tibble)
85 library(stats)
86 library(dplyr)
87 library(ggplot2)
88
89 ## Create data to plot
90 Conc=seq(0, 200, 0.1)
91 G1.pred=predict(G1.glm, newdata=data.frame(Concentration=Conc),
92               type="response")
93 Conc=seq(0, 200, 0.1)
94 G2.pred=predict(G2.glm, newdata=data.frame(Concentration=Conc),
95               type="response")
96 Conc=seq(0, 200, 0.1)
97 G3.pred=predict(G3.glm, newdata=data.frame(Concentration=Conc),
98               type="response")
99
100 ## some data to predict at: 100 values over the range of leafHeight
101 G1.ndata <- data.frame(Concentration=Conc, Mortality=G1.pred)
102 G2.ndata <- data.frame(Concentration=Conc, Mortality=G2.pred)
103 G3.ndata <- data.frame(Concentration=Conc, Mortality=G3.pred)
104

```

Esta parte do script calcula os intervalos de confiança ajustados aos parâmetros do modelo criado para os valores calculados na etapa anterior.

```

105 ## grad the inverse link function
106 ilink1l <- family(G1.glm)$linkinv
107 ## add fit and se.fit on the **link** scale
108 G1.ndata <- bind_cols(G1.ndata,
109                     setNames(as_tibble(predict(G1.glm, G1.ndata, se.fit = TRUE)[1:2]),
110                               c('fit_link','se_link'))))
111 ## create the interval and backtransform
112 G1.ndata <- mutate(G1.ndata,
113                   fit_resp = ilink1l(fit_link),
114                   right_upr = ilink1l(fit_link + (2 * se_link)),
115                   right_lwr = ilink1l(fit_link - (2 * se_link)))
116

```

O primeiro tipo de gráfico é uma curva concentração-mortalidade que inclui um sombreado representando intervalos de confiança ajustados ao longo da curva. Nesta parte do script existem muitos parâmetros como tamanho da letra ou cor usada (em azul/"verde") e detalhes como títulos (em "verde") que podem ser ajustados pelo utilizador.

```
141 ##Plot type 1
142 G1.sg <- ggplot(G1.ndata, aes(x = Concentration, y = Mortality)) +
143   scale_x_continuous(limits = c(0,200), expand = expansion(mult = c(0, .02))) +
144   scale_y_continuous(limits = c(0,1), expand = expansion(mult = c(0, .02))) +
145   geom_line(col="steelblue", size=1) +
146   geom_point(data=G1, shape=16, aes(x = Concentration, y = Mortality), col="steelblue", size=2)
147   labs(x = 'Concentration (mg/L)', y = 'Mortality', title = "a. Aedes aegypti - Pop01") +
148   geom_segment(aes(x = 0, y = 0.50, xend = G1.lc [1], yend = 0.50),
149     col="steelblue", size=1, linetype = "dashed") +
150   geom_segment(aes(x = G1.lc [1], y = 0, xend = G1.lc [1], yend = 0.50),
151     col="steelblue", size=1, linetype = "dashed") +
152   geom_segment(aes(x = 0, y = 0.90, xend = G1.lc [2], yend = 0.90),
153     col="steelblue", size=1, linetype = "dashed") +
154   geom_segment(aes(x = G1.lc [2], y = 0, xend = G1.lc [2], yend = 0.90),
155     col="steelblue", size=1, linetype = "dashed") +
156   theme(plot.title = element_text(size = 10, color = "black")) +
157   theme(axis.title.y = element_text(size = 8, angle = 90))+
158   theme(axis.title.x = element_text(size = 8, angle = 00)) +
159   theme_bw()
160 G1.sg <- G1.sg + geom_ribbon(data = G1.ndata, aes(ymin = right_lwr, ymax = right_upr),
161   alpha = 0.1, fill="steelblue", lty = 3, col = "steelblue")
```

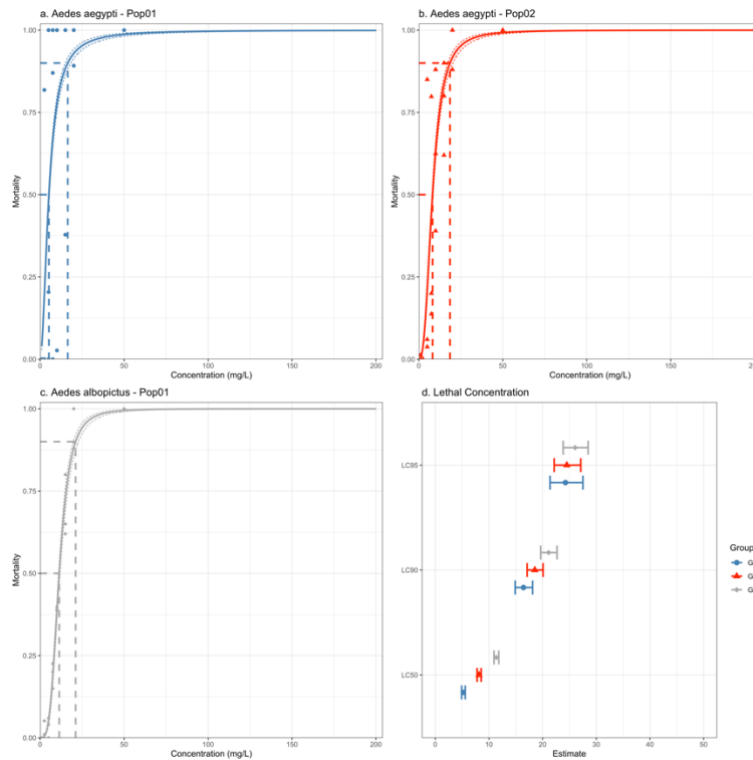
O segundo tipo de gráfico permite comparar a variação das concentrações letais em múltiplas populações. As linhas de código entre as linhas 209-218 são necessárias para compilar os dados de cálculo das concentrações letais "G*.lc" de forma a serem usados pelo "ggplot".

```
208 ##Plot type 2
209 G1.dflc<-as.data.frame(G1.lc)
210 G1.dflc$Group<-c("G1", "G1", "G1")
211 G1.dflc$Type<-c("LC50", "LC90", "LC95")
212 G2.dflc<-as.data.frame(G2.lc)
213 G2.dflc$Group<-c("G2", "G2", "G2")
214 G2.dflc$Type<-c("LC50", "LC90", "LC95")
215 G3.dflc<-as.data.frame(G3.lc)
216 G3.dflc$Group<-c("G3", "G3", "G3")
217 G3.dflc$Type<-c("LC50", "LC90", "LC95")
218 LC.dflc<-rbind(G1.dflc, G2.dflc, G3.dflc)
219 LC.graph<- ggplot(data=LC.dflc, aes(x=Type, y=LC, colour = Group, group = Group, shape = Group)) +
220   geom_point(size = 3, position = position_dodge(width = 0.5))+
221   scale_color_manual(values = c("steelblue", "red", "darkgray")) +
222   scale_shape_manual(values = c(16,17,18)) +
223   geom_errorbar(aes(ymin = Lower, ymax = Upper), position = position_dodge(width = 0.5), size=1, width=.4) +
224   coord_flip() + # flip coordinates (puts labels on y axis)
225   # geom_hline(yintercept=100, lty=2) +
226   xlim("LC50", "LC90", "LC95") + ylim(0, 50) + # add a dotted line at x=1 after flip
227   labs(x= " ", y = 'Estimate', title = "d. Lethal Concentration") +
228   theme_bw() # use a white background
229 LC.graph
```

Na preparação de artigos e relatórios é importante a qualidade e formato das imagens. O R permite exportar os gráficos de várias formas, mas nós recomendamos o uso do pacote "gridExtra" compilar múltiplos gráficos e uso o "ggsave" para definir os critérios para criar a imagem com os parâmetros necessários a alta qualidade necessários para impressão.

```
231 ##Export figure with multiple graphics
232 library(gridExtra)
233 Graph1 <- arrangeGrob(G1.sg, G2.sg, G3.sg, LC.graph, nrow=2) #generates g
234 ggsave("figure1.png", Graph1, units="mm", width=340, height=340, dpi=600)
235
```

Imagem final gerada é organizada em quatro quadrantes com gráficos individualizados com a curva de concentração-resposta com destaque da localização das concentrações letais a 50% e 90%, enquanto o quarto gráfico mostra uma comparação entre as concentrações letais das três populações que inclui os intervalos de confiança a 95%.



3.7 Questões e comentários

- Tradicionalmente informação sobre “slope” e outros parâmetros do modelo são importantes de obter. Para obter esta informação escreva na *Console* – `summary(G1.glm)`. Esta função vai permitir ver o sumário do modelo onde pode ser obtido o “slope” no coeficiente designado `log(Concentration)` e outros parâmetros sobre a qualidade do modelo como é o caso dos valores de desvios e o AIC (*Akaike information criterion*).

```

Console Terminal Background Jobs
R 4.2.1 ~ /Documents/R-Tutorial_IC/
> summary(G1.glm)

Call:
glm(formula = cbind(Dead, Alive) ~ log(Concentration), family = binomial(link = "logit"),
    data = G1)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-14.185   -5.655    1.505    4.980   13.357

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -3.17263    0.16574  -19.14  <2e-16 ***
log(Concentration)  1.91824    0.08227   23.32  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2463.1 on 23  degrees of freedom
Residual deviance: 1313.1 on 22  degrees of freedom
AIC: 1341.7

Number of Fisher Scoring iterations: 6
>

```

- Como variam as concentrações letais da substância “S2” em relação aos nossos resultados com “S1”?

Altera os dados de 3.3 de forma a realizar a mesma análise até 3.5 com uma substância diferente

- Podemos criar os gráficos para “S2” sem ajustar o nosso script?

Corram o ponto 3.6 sem alterações para verificar as diferenças e identificar parâmetros que tenham de ser alterados (exemplos: concentrações, títulos, limites de eixos, cores).

- Consegue analisar múltiplas substâncias para a mesma população?

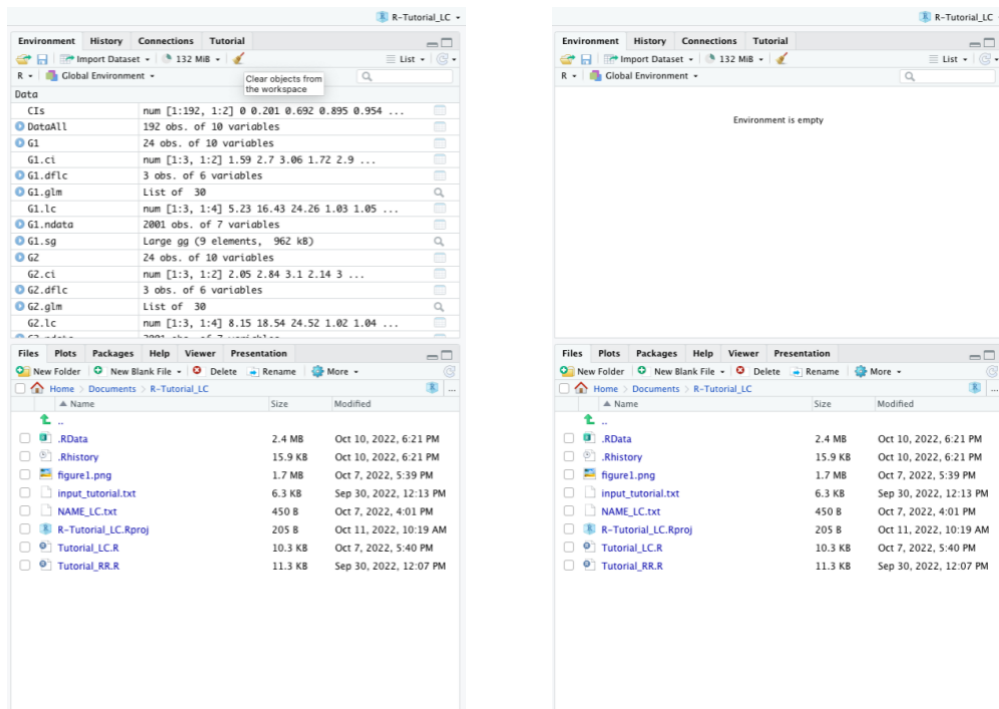
O ficheiro input só tem duas substâncias definidas

4. Abordagem 2: cálculo de rácio de resistência com linhagem de referência

Nesta segunda abordagem temos de abrir o segundo *Script* designado *Tutorial_RR.R*. A parte inicial deste *Script* é equivalente ao *Tutorial_LC.R* (pontos 3.1 a 3.5) tendo apenas uma pequena diferença na natureza dos subgrupos definidos (equivalente ao ponto 3.3).

Antes de iniciar a análise do novo *Script* é importante verificar que os recursos usados nas análises anteriores foram apagados para garantir que não tenham influência na nova análise.

- No *Environment*, o símbolo da “vassoura” permite apagar todos os recursos na memória de trabalho.



Na definição dos subgrupos, o primeiro subgrupo (G1Ref) é considerado a referência para os subsequentes cálculos no *Script*. Este subgrupo tem de obrigatoriamente conter os dados de linhagens laboratoriais consideradas susceptíveis para inseticidas químicos neurotóxicos, como é o caso da linhagem Rockefeller, para permitir calcular o “Rácio de Resistência”. Os outros subgrupos devem ser utilizados para outras populações de mosquito. Este tipo de análise implica o uso da mesma substância para todos os subgrupos para os cálculos serem validos.

```
27 ## create multiple data sets
28 G1Ref=DataAll[DataAll$Species=="Aeegypti" &
29             DataAll$Lineage=="Rockefeller" &
30             DataAll$Substance=="S1" ,]
31 G2=DataAll[DataAll$Species=="Aeegypti" &
32            DataAll$Lineage=="Pop01" &
33            DataAll$Substance=="S1" ,]
34 G3=DataAll[DataAll$Species=="Aeegypti" &
35            DataAll$Lineage=="Pop02" &
36            DataAll$Substance=="S1" ,]
37
```

4.1 Determinar Rácio de Resistência

Após o cálculo das concentrações letais para as três linhagens, nós iremos utilizar a função *MOVERR* do pacote “*pairwiseCI*” para determinar o “Rácio de Resistência” com intervalos de

confiança a 95% para as duas populações. Esta parte do *Script* é estruturalmente estável para uma análise onde se calcula o “Rácio de Resistência” para duas populações em relação a uma linhagem, a partir de três estimativas distintas de concentração letal. Caso deseje alterar o número de subgrupos ou estimativas, os parâmetros de localização dos valores de concentração letal e intervalos de confiança na matriz de cálculo ($G^*.lc$) têm de ser ajustados.

```
69 ##Calculate RRs for G2 and G3 with G1Ref as reference - Install pairwiseCI
70 library(pairwiseCI)
71 #RR for G2
72 G2.RR50<- MOVErr(theta0 = G1.lc[1], ci0 = c(G1.lc[7], G1.lc[10]),
73               theta1 = G2.lc[1], ci1 = c(G2.lc[7], G2.lc[10]),
74               alternative = "two.sided")
75 G2.RR50est<-do.call(rbind.data.frame, G2.RR50[2])
76 G2.RR50ci<-do.call(rbind.data.frame, G2.RR50[1])
77 G2.RR50data<-cbind(G2.RR50est[1],G2.RR50ci[1],G2.RR50ci[2])
78 dimnames(G2.RR50data)[[2]] <- c("Ratio", "Lower", "Upper")
79
```

Após calcular o “Rácio de Resistência” para as duas populações, os resultados são exportados em um ficheiro de texto para manter como registro. O nome do ficheiro pode ser ajustado conforme necessário.

```
134 RR_final<-rbind(G2.RRtable, G3.RRtable)
135 sink("NAME_RR.txt")
136 print(RR_final)
137 sink()
```

4.2 Gráficos

A preparação de gráficos pode ser dividida em duas etapas: 1) a organização de dados usados para criar os gráficos, 2) os comandos de código para criar a imagem gráfica e exportá-la. A organização de dados neste *Script* é idêntica ao *Tutorial_LC.R* não tendo variações.

Na etapa de criação de gráficos, as linhas de comando variam em relação ao *Script* anterior mesmo que usem as mesmas ferramentas. Várias alterações são incluídas no *Script* de forma a maximizar a comparação entre as populações:

- Os gráficos iniciais são de escala semilogarítmica com o eixo do x apresentando uma escala de logaritmo base 10.
- Os gráficos contêm informação de uma população (cor varia entre população) e da linhagem de referência (cor azul claro).
- Os gráficos não contêm linhas para destacar as concentrações letais 50% e 90%.

```
198 ##Plot type 1
199 G2.graph <- ggplot(G1.ndata, aes(x = Concentration, y = Mortality)) +
200   scale_x_continuous(trans='log10', expand = expansion(mult = c(0, .02))) +
201   scale_y_continuous(limits = c(0,1), expand = expansion(mult = c(0, .02))) +
202   geom_line(col="steelblue", size=1) +
203   geom_point(data=G1Ref, shape=16, aes(x = Concentration, y = Mortality), col="steelblue", size=2) +
204   geom_line(data=G2.ndata, aes(x = Concentration, y = Mortality), col="red", size=1) +
205   geom_point(data=G2, shape=17, aes(x = Concentration, y = Mortality), col="red", size=2) +
206   labs(x = 'Concentration (mg/L)', y = 'Mortality', title = "a. Aedes aegypti - Pop1") +
207   theme(plot.title = element_text(size = 10, color = "black")) +
208   theme(axis.title.y = element_text(size = 8, angle = 90)) +
209   theme(axis.title.x = element_text(size = 8, angle = 0)) +
210   theme_bw()
211 G2.graph <- G2.graph + geom_ribbon(data = G1.ndata, aes(ymin = right_lwr, ymax = right_upr),
212                                alpha = 0.1, fill="steelblue", lty = 3, col = "steelblue") +
213   geom_ribbon(data = G2.ndata, aes(ymin = right_lwr, ymax = right_upr),
214             alpha = 0.1, fill="red", lty = 3, col="red")
215 G2.graph
```

- O gráfico final apresenta informação sobre o “Rácio de Resistência” e seus intervalos de confiança para duas populações. Incluindo os limites designados pela OMS para classificar uma população suscetível/resistente para determinado inseticida.

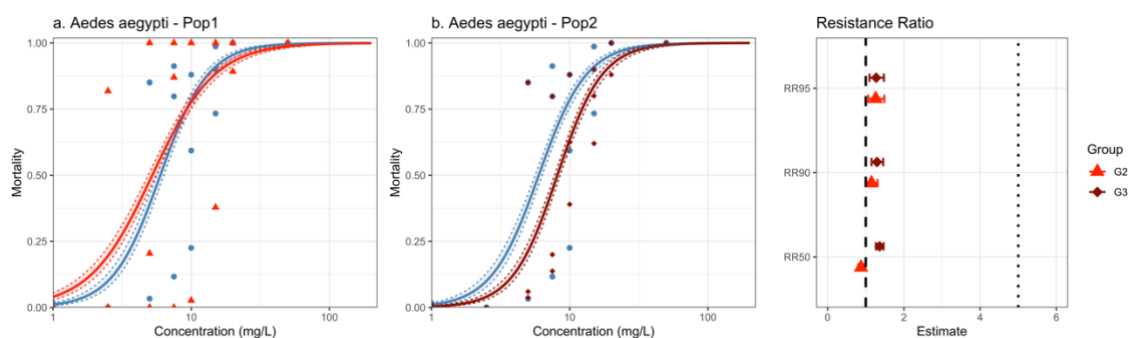
- A organização dos três gráficos numa única linha foi ajustada para criar uma imagem para o tutorial.

```

235 ##Plot type 2
236 RR.graph<- ggplot(data=RR_final, aes(x=Type, y=Ratio, colour = Group, group = Group)) +
237   geom_point(size = 4, position = position_dodge(width = 0.5))+
238   scale_color_manual(values = c("red", "darkred")) +
239   scale_shape_manual(values = c(17,18)) +
240   geom_errorbar(aes(ymin = Lower, ymax = Upper), position = position_dodge(width = 0.5), size=.8, width=.2) +
241   coord_flip() + # flip coordinates (puts labels on y axis)
242   geom_hline(yintercept=1, lty=2, lwd=1) +
243   geom_hline(yintercept=5, lty=3, lwd=1) +
244   xlim("RR50", "RR90", "RR95") + ylim(0, 10) + # add a dotted line at x=1 after flip
245   labs(x= " ", y = 'Estimate', title = "Resistance Ratio") +
246   theme_bw() # use a white background
247 RR.graph

```

Alguns parâmetros para o uso as funções do pacote “ggplot2” podem ser ajustadas da mesma forma que previamente treinado no *Tutorial_LC.R*. A informação de apoio do pacote “ggplot2” também contém muita informação útil para o uso de representações visuais.



4.3 Questões e comentários

- Como variam os rácios de resistência da substância “S2” em relação aos nossos resultados com “S1”?

- Podemos criar os gráficos para “S2” sem ajustar o nosso script?

- **DESAFIO:** Consegue alterar o *Script* para representar os “Rácios de Resistência” para duas substâncias em simultâneo?

Este desafio têm o objetivo o utilizador alterar de forma significativa o Script e pode ser resolvido de múltiplas formas.