**Solution to Exercise 1: Introduction to R software: basics**

**Key points:**

- a. R is both a language and an environment of computing
- b. Anything and everything can be assigned to an object
- c. Object names are case-sensitive
- d. A vector is a one-dimensional object of like elements
- e. A matrix is a two-dimensional table (another object) of like elements
- f. Vectors can be bound to matrices

---

**Task**

*In analogy to the calculation of the odds ratio, write an R script `e_ex01_rr.r` that calculates the relative risk from two proportions (thus not ratio of incidence rates, but ratio of two prevalence proportions) for a study used in Altman’s textbook. It has the isolation of *Helicobacter pylori* as the outcome and the history of an ulcer in the mother as the exposure. We use the following set-up of notations:*

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Outcome</th>
<th>Ill</th>
<th>Healthy</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>A</td>
<td>B</td>
<td>A+B</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>C</td>
<td>D</td>
<td>C+D</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>A+B</td>
<td>B+D</td>
<td>A+B+C+D</td>
<td></td>
</tr>
</tbody>
</table>

The data provided by Altman in table 7.2 (page 59) are as follows:

<table>
<thead>
<tr>
<th>Mother with a history of ulcer</th>
<th><em>H pylori</em> isolated</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>No</td>
<td>112</td>
<td>729</td>
</tr>
<tr>
<td>Total</td>
<td>118</td>
<td>745</td>
</tr>
</tbody>
</table>

The relative risk is calculated by:

\[ R = \frac{(A/ (A+B) )}{(C/ (C+D) )} \]
The standard error of the logeR is:
\[ se = \sqrt{\frac{1}{A} - \frac{1}{A+B} + \frac{1}{C} - \frac{1}{C+D}} \]

The 95% CI is thus:
\[ 95\% CI = \exp(\text{logeR} \pm 1.96 \times se) \]

Solution:
The solution for the e_ex01_rr.r is a straight forward modification of the e_ex01_or.r:

```r
# Complete 2 by 2 table from A, B, C, D and calculate relative risk with Wolf CI
A <- 54
B <- 89
C <- 60
D <- 245
AB <- c(A, B, A+B)
CD <- c(C, D, C+D)
ABCD <- rbind(AB, CD)
rtot <- abcd[1,]+abcd[2,]
tab <- rbind(ABCD, rtot)
colnames(tab) <- cbind("Ill", "Healthy", "Total")
rownames(tab) <- c("Exp+", "Exp-", "Total")
names(dimnames(tab)) <- c("Exposure", "Status")
se <- sqrt(1/A - 1/(A+B) + 1/C - 1/(C+D))
rr <- (tab[1,1]/tab[1,3])/(tab[2,1]/tab[2,3]); rr
rr.ci.lower <- exp(log(rr)-1.96*se)
rr.ci.upper <- exp(log(rr)+1.96*se)

print(tab)
cat("\nRR: ", round(rr, digits=3), " \n95% CI: ", round(rr.ci.lower, digits=3), round(rr.ci.upper, digits=3), "\n")
```

The result is:
```
     Ill Healthy Total
Exp+  6   16  22
Exp- 112  729 841
Total 118  745 863

RR: 2.048
95% CI: 1.013 4.14
```