Exercise 1 (Cohort study)

The table shows data from a cohort study representing the prevalence of hay fever and eczema in 11 year old children:

	Hay fever yes	Hay fever no	Total
Eczema yes	141	420	561
Eczema no	928	13 525	14 463
Total	1069	13 945	15 014

(a) Calculate the risk ratio (RR) and odds ratio (OR) for hay fever in people with Eczema compared to people without. Interpret the results. (**Hint**: Start by defining the variables a, b, c ,d as in the lecture slides, page 10. Then apply the formulas.)

(b) Calculate the risk ratio (RR) and odds ratio (OR) for Eczema in people with Hay fever compared to people without. Interpret the results. Do you notice anything interesting when you compare the results to the results of the previous subtask?

(c) Is there evidence for an association between Eczema and Hay fever? Choose an appropriate test. (**R-Hint**: `chisq.test()`. The function expects a cross table.)

Exercise 2 (Case control study)

We are using the data from a case control study investigating the association between coffee drinking and pancreatic cancer. The dataset `coffee.csv` can be downloaded from the website. You can read in the data using `read.csv(..., sep=";", header=TRUE)`.

(a) In general it is not possible to calculate the Risk Ratio from a Case control study. Why?

(b) To investigate your dataset, calculate a contingency table for coffee drinking vs. pancreatic cancer. Additionally, show the frequencies graphically. What do you observe? (**R-Hint**: `table()`, `mosaicplot()`)

(c) Compute the OR and the corresponding 95% CI with the function `fisher.test()` measuring the association between coffee drinking and pancreatic cancer. Intrepret the result. Is there evidence for an association between coffee drinking and pancreatic cancer?