

# EpiStats

*jean.pierre.decorps@gmail.com*

*2020-04-15*

---

# Package EpiStats

## Description

The EpiStats package is a set of functions aimed at epidemiologists. They include commands for measures of association and impact for case control studies and cohort studies. They may be particularly useful for outbreak investigations and include univariate and stratified analyses.

The generic function *crossTable* provides a contingency table with optional parameters *percent* and *statistic*

The functions for cohort studies include the *CS*, *CSTable* and *CSInter* commands.

The functions for case control studies include the *CC*, *CCTable* and *CCInter* commands.

All variables used need to be numeric binary variables and coded as 0 and 1 or as factors.

## Cohort study functions:

The cohort study functions relate to cohort studies that measure risks, rather than rates in person- time.

The **CS** function provides a 2 by 2 table and measures the association between the outcome and one exposure. It includes the risk ratio and its 95% confidence intervals, the attributable fraction among the exposed and unexposed, and a chi square test and its p-value.

The **CSTable** function displays the measures of association between the outcome and a set of exposures in a table (risk ratios, confidence intervals and p-values). This helps the researcher to compare between exposures and provides a nice table for reports.

The **CSInter** function investigates the effect of a third variable on the association between an exposure and the outcome. It presents two by two tables stratified by the levels of a third value. It provides the Woolf test for homogeneity between stratum-specific risk ratios. It provides the crude risk ratio between an exposure and an outcome and the risk ratio adjusted by the third variable. *CSInter* helps the researcher understand whether a third variable may have an effect modifying or confounding effect on the association between an exposure and the outcome.

## Case control study functions:

The **CC** function provides a 2 by 2 table and measures the association between the outcome and one exposure. It includes the odds ratio and its 95% confidence intervals, the attributable fraction among the exposed, and a chi square test and its p-value.

The **CCTable** function displays the measures of association between the outcome and a set of exposures in a table (odds ratios, confidence intervals and p-values). This helps the researcher to compare between exposures and provides a nice table for reports.

The **CCInter** function investigates the effect of a third variable on the association between an exposure and the outcome. It presents two by two tables stratified by the levels of a third value. It provides the Woolf test for homogeneity between stratum-specific odds ratios. It provides the crude odds ratio between an exposure and an outcome and the odds ratio adjusted by the third variable. *CCInter* helps the researcher understand whether a third variable may have an effect modifying or confounding effect on the association between an exposure and the outcome.

## The “Tiramisu” dataset

The dataset used in this vignette is from an outbreak investigation carried out in Germany in 1998 by Anja Hauri, Robert Koch Institute. It is used in case studies by organisations including EPIET, ECDC and EpiConcept.

---

The **CSTable**, **CSInter**, **CCTable** and **CCInter** functions are based on commands written in Stata by *Gilles Desve*, who we gratefully acknowledge.

---

# Working with EpiStats and “Tiramisu” dataset

## Loading and recoding the dataset

```
library(EpiStats)
library(dplyr)
library(knitr)

options(knitr.kable.NA = '')
#options(width=200)

data(Tiramisu)
DF <- Tiramisu

DF <- DF %>%
  # filter(age != "NA") %>%
  mutate(agegroup = case_when(age < 30 ~ 0, age >= 30 ~ 1)) %>%
  mutate(tportion = case_when(tportion == 0 ~ 0, tpotion == 1 ~ 1, tpotion >= 2 ~ 2)) %>%
  mutate(tportion = as.factor(tportion)) %>%
  as.data.frame(stringsAsFactors=TRUE)

Colnames <- DF %>%
  select(-ill, -age, -dateonset, -uniquekey, -tportion, -mportion) %>%
  colnames()
```

## crossTable

Creates a contingency table of variable of interest and exposure. Percentage are optionals by *row* or by *column*. It can provides an optional statistic (*fisher* or *chisquare*).

### Syntax

```
crossTable(data, var1, var2, percent="none", statistic="none")
```

### Examples

Recoding some data to have ordered factors

```
DF2 <- DF
DF2$ill <- factor(DF2$ill, levels=c(1,0), ordered = TRUE)
DF2$beer <- factor(DF2$beer, levels=c(1,0), ordered = TRUE)
DF2$tira <- factor(DF2$tira, levels=c(1,0), ordered = TRUE)
DF2$sex <- factor(DF2$sex, levels = c("males", "females"), ordered = TRUE)
```

#### Example 1: crossTable ill - tira

```
ret <- crossTable(DF2, var1="ill", var2="tira")
ret
```

```
##  tira / ill   1   0 Total
## 1         1  94  27  121
## 2         0   7 158  165
## 3      Total 101 185  286
```

```
kable(ret, align="r")
```

tira / ill	1	0	Total
1	94	27	121
0	7	158	165
Total	101	185	286

#### Example 2: crossTable ill - sex with column percentage and chi2 stat

```
ret <- crossTable(DF2, "ill", "sex", "col", "chi2")
kable(ret, align="r", caption = "with columns %")
```

Table 2: with columns %

sex / ill	1	0	Total
males	50	102	152
%	48.54	54.26	52.23
females	53	86	139
%	51.46	45.74	47.77
Total	103	188	291
%	100.00	100.00	100.00
-	-	-	-
Pearson CHI2	0.8701	Pr	0.351

### Example 3: CrossTable ill - sex with row percentage and Fisher stat

NB: All parameters are unquoted

```
ret <- crossTable(DF2, ill, sex, row, fisher)
ret
```

```
##      sex / ill      1      %      0      % Total      %
## 1      males      50 32.89 102 67.11   152 100
## 2      females     53 38.13   86 61.87   139 100
## 3      Total     103 35.40 188 64.60   291 100
## 4      -          -      -      -      -      -      -
## 5 Fisher's exact 0.391      -      -      -      -      -
```

```
kable(ret, align="r")
```

sex / ill	1	%	0	%	Total	%
males	50	32.89	102	67.11	152	100
females	53	38.13	86	61.87	139	100
Total	103	35.40	188	64.60	291	100
-	-	-	-	-	-	-
Fisher's exact	0.391	-	-	-	-	-

### CrossTable beer - sex with column and row percentages and Chi2 stat

NB: All parameters are unquoted

```
ret <- crossTable(DF2, beer, sex, both, chi2)
ret
```

```
##      sex / beer      1      %      0      % Total      %
## 1      males      84 59.15      58 40.85   142 100.00
## 2      %      79.25      - 35.15      -      -      -
## 3      females     22 17.05     107 82.95   129 100.00
## 4      %      20.75      - 64.85      -      -      -
## 5      Total     106 39.11     165 60.89   271 100.00
## 6      %     100.00      - 100.00      - 100.00      -
## 7      -          -      -      -      -      -      -
## 8 Pearson CHI2 50.3078      Pr      0      -      -      -
```

```
kable(ret, align="r", caption = "% rows and columns")
```

Table 4: % rows and columns

sex / beer	1	%	0	%	Total	%
males	84	59.15	58	40.85	142	100.00
%	79.25	-	35.15	-	-	-
females	22	17.05	107	82.95	129	100.00
%	20.75	-	64.85	-	-	-
Total	106	39.11	165	60.89	271	100.00
%	100.00	-	100.00	-	100.00	-
-	-	-	-	-	-	-
Pearson CHI2	50.3078	Pr	0	-	-	-

## CS

CS analyses cohort studies with equal follow-up time per subject. The risk (the proportion of individuals who become cases) is calculated overall and among the exposed and unexposed. Note that all variables need to be numeric and binary and coded as "0" and "1".

Point estimates and confidence intervals for the risk ratio and risk difference are calculated, along with attributable or preventive fractions for the exposed and the total population. Additionally you can select if you want to display the Fisher's exact test, by specifying `exact = TRUE`. If you specify `full = TRUE` you can easily access useful statistics from the output tables.

### Syntax

```
CS(x, cases, exposure, exact, full=FALSE)
```

### Example 1: CS ill - mousse (unformatted)

```
CS(DF, "ill", "mousse", exact = FALSE)
```

```
## $df1
##           Cases Non Cases Total Risk
## Exposed      81      42   123 0.66
## Unexposed    22     144   166 0.13
## Total       103     186   289 0.36
##
## $df2
##           Point estimate 95%CI ll 95%CI ul
## Risk difference           0.53   0.43   0.62
## Risk ratio                4.97   3.30   7.48
## Attr. frac. ex.           0.80   0.70   0.87
## Attr. frac. pop           0.63    NA    NA
## chi2(1)                   85.22    NA    NA
## Pr>chi2                    0.000    NA    NA
```

## Example 2: CS ill - beer (formatted)

The following results tables are outputs in “markdown” using the *kable* function.

```
result <- CS(DF, "ill", "beer", exact = TRUE, full = TRUE)
kable(result$df1, align = "r")
```

	Cases	Non Cases	Total	Risk
Exposed	30	76	106	0.28
Unexposed	69	96	165	0.42
Total	99	172	271	0.37

```
kable(result$df2, align = result$df2.align )
```

	Point estimate	95%CI ll	95%CI ul
Risk difference	-0.14	-0.25	-0.02
Risk ratio	0.68	0.48	0.96
Prev. frac. ex.	0.32	0.04	0.52
Prev. frac. pop	0.13	NA	NA
chi2(1)	5.09	NA	NA
Pr>chi2	0.024	NA	NA
Fisher p.value	0.028	NA	NA

By storing the results in the object “result”, you are able to use the result tables in Markdown as shown above. By specifying “full = TRUE” you can also easily use individual elements of the results. For example if you would like to view just the risk ratio, you can view it by typing:

```
result$st$risk_ratio$point_estimate
```

```
## [1] 0.6767842
```

## CSTable - Summary table for cohort studies

CSTable is used for univariate analysis of cohort studies with several exposures. The results are summarised in one table with one row per exposure making comparisons between exposures easier and providing a useful table for integrating into reports. Note that all variables need to be numeric and binary and coded as “0” and “1”.

The results of this function contain: The name of exposure variables, the total number of exposed, the number of exposed cases, the attack rate among the exposed, the total number of unexposed, the number of unexposed cases, the attack rate among the unexposed, risk ratios, 95% confidence intervals, 95% p-values.

You can optionally choose to display the Fisher’s exact p-value instead of the Chi squared p-value, with the option `exact = TRUE`.

You can specify the sort order, with the option `sort=“rr”` to order by risk ratios. The default sort order is by p-values.

The option `“full = TRUE”` provides you with useful formatting information, which can be handy if you’re using “markdown”.

### Syntax

```
CSTable(x, cases, exposure=c(), exact=FALSE, sort = “pvalue”, full=FALSE)
```

### Example 1: CSTable results ordered by p-value (unformatted)

```
CSTable(DF,
  "ill",
  exposure = c("sex", "agegroup", "tira", "beer", "mousse", "wmousse", "dmousse",
    "redjelly", "fruitsalad", "tomato", "mince", "salmon", "horseradish",
    "chickenwin", "roastbeef", "pork"))
```

```
## $df
##          Tot.Exp. Exp.Cases  AR% Tot.Unex. Unex.Cases  AR%  RR
## tira          121         94 77.69      165          7  4.24 18.31
## mousse         123         81 65.85      166         22 13.25  4.97
## wmousse         72         49 68.06      205         49 23.90  2.85
## dmousse        113         76 67.26      174         26 14.94  4.50
## redjelly        79         45 56.96      212         58 27.36  2.08
## fruitsalad      71         46 64.79      220         57 25.91  2.50
## beer           106         30 28.30      165         69 41.82  0.68
## tomato          83         35 42.17      208         68 32.69  1.29
## pork           120         48 40.00      169         54 31.95  1.25
## horseradish     72         30 41.67      217         72 33.18  1.26
## sex            152         50 32.89      139         53 38.13  0.86
## roastbeef       29          8 27.59      262         95 36.26  0.76
## chickenwin     84         33 39.29      207         70 33.82  1.16
## mince           87         32 36.78      204         71 34.80  1.06
## agegroup        68         25 36.76      215         75 34.88  1.05
## salmon         104         37 35.58      183         63 34.43  1.03
##          CI ll CI ul p(Chi2)
## tira          8.81 38.04  0.000
## mousse         3.30  7.48  0.000
## wmousse        2.13  3.81  0.000
## dmousse        3.09  6.56  0.000
## redjelly       1.56  2.79  0.000
## fruitsalad     1.89  3.31  0.000
## beer           0.48  0.96  0.024
## tomato         0.94  1.77  0.127
## pork           0.92  1.71  0.158
## horseradish    0.90  1.75  0.192
## sex            0.63  1.18  0.351
## roastbeef     0.41  1.40  0.354
```



## chickenwin	0.84	1.61	0.377
## mince	0.76	1.48	0.747
## agegroup	0.73	1.51	0.777
## salmon	0.75	1.43	0.844

## Example 2: CStable results ordered by risk ratio (formatted)

The following results tables are outputs in “markdown” using the kable function.

```
res = CStable(DF, "ill", sort = "rr", exposure = Colnames, full = TRUE)
kable(res$df, digits=res$digits, align=res$align)
```

	Tot.Exp.	Exp.Cases	AR%	Tot.Unex.	Unex.Cases	AR%	RR	CI ll	CI ul	p(Chi2)
tira	121	94	77.69	165	7	4.24	18.31	8.81	38.04	0.000
mousse	123	81	65.85	166	22	13.25	4.97	3.30	7.48	0.000
dmousse	113	76	67.26	174	26	14.94	4.50	3.09	6.56	0.000
wmousse	72	49	68.06	205	49	23.90	2.85	2.13	3.81	0.000
fruitsalad	71	46	64.79	220	57	25.91	2.50	1.89	3.31	0.000
redjelly	79	45	56.96	212	58	27.36	2.08	1.56	2.79	0.000
tomato	83	35	42.17	208	68	32.69	1.29	0.94	1.77	0.127
horseradish	72	30	41.67	217	72	33.18	1.26	0.90	1.75	0.192
pork	120	48	40.00	169	54	31.95	1.25	0.92	1.71	0.158
chickenwin	84	33	39.29	207	70	33.82	1.16	0.84	1.61	0.377
mince	87	32	36.78	204	71	34.80	1.06	0.76	1.48	0.747
agegroup	68	25	36.76	215	75	34.88	1.05	0.73	1.51	0.777
salmon	104	37	35.58	183	63	34.43	1.03	0.75	1.43	0.844
sex	152	50	32.89	139	53	38.13	0.86	0.63	1.18	0.351
roastbeef	29	8	27.59	262	95	36.26	0.76	0.41	1.40	0.354
beer	106	30	28.30	165	69	41.82	0.68	0.48	0.96	0.024

### Example 3: CStable results ordered by p-value from the Fisher's exact test (formatted)

The following results tables are outputs in "markdown" using the kable function.

```
res = CStable(DF, "ill", exact = TRUE, exposure = Colnames, full = TRUE)
kable(res$df, digits=res$digits, align=res$align)
```

	Tot.Exp.	Exp.Cases	AR%	Tot.Unex.	Unex.Cases	AR%	RR	CI ll	CI ul	p(Fisher)
tira	121	94	77.69	165	7	4.24	18.31	8.81	38.04	0.000
wmousse	72	49	68.06	205	49	23.90	2.85	2.13	3.81	0.000
dmousse	113	76	67.26	174	26	14.94	4.50	3.09	6.56	0.000
mousse	123	81	65.85	166	22	13.25	4.97	3.30	7.48	0.000
redjelly	79	45	56.96	212	58	27.36	2.08	1.56	2.79	0.000
fruitsalad	71	46	64.79	220	57	25.91	2.50	1.89	3.31	0.000
beer	106	30	28.30	165	69	41.82	0.68	0.48	0.96	0.028
tomato	83	35	42.17	208	68	32.69	1.29	0.94	1.77	0.137
pork	120	48	40.00	169	54	31.95	1.25	0.92	1.71	0.171
horseradish	72	30	41.67	217	72	33.18	1.26	0.90	1.75	0.203
sex	152	50	32.89	139	53	38.13	0.86	0.63	1.18	0.391
roastbeef	29	8	27.59	262	95	36.26	0.76	0.41	1.40	0.417
chickenwin	84	33	39.29	207	70	33.82	1.16	0.84	1.61	0.418
agegroup	68	25	36.76	215	75	34.88	1.05	0.73	1.51	0.773
mince	87	32	36.78	204	71	34.80	1.06	0.76	1.48	0.789
salmon	104	37	35.58	183	63	34.43	1.03	0.75	1.43	0.898

By storing the results in the object "res", you are able to use the result table in Markdown as shown above. You can also use individual elements of the results. For example if you would like to view just the risk ratio, you can view it by typing (for example):

```
res$df$'Risk Ratio'[2]
```

```
## NULL
```

## CSInter - Stratified analysis for cohort studies

CSInter is useful to determine the effects of a third variable on the association between an exposure and an outcome. CSInter produces 2 by 2 tables with stratum specific risk ratios, attributable risk among exposed and population attributable risk. Note that the outcome and exposure variable need to be numeric and binary and coded as "0" and 1". The third variable needs to be numeric, but may have more categories, such as "0", "1" and "2".

CSInter displays a summary with the crude RR, the Mantel Haenszel adjusted RR and the result of a "Woolf" test for homogeneity of stratum-specific RR.

The option "full = TRUE" provides you with useful formatting information, which can be handy if you're using "markdown".

### Syntax

```
CSInter(x, cases, exposure, by, full=FALSE)
```

### Example 1 : CSInter ill - wmouse by tira (unformatted)

```
CSInter(DF, cases="ill", exposure = "wmousse", by = "tira")
```

```
## $df1
##   CSInter ill - wmouse by(tira) Total Cases Risk %          P.est. Stats
## 1                tira = 1   112 <NA>      NA Risk difference  0.06
## 2                Exposed    52   43  82.69      Risk Ratio   1.08
## 3                Unexposed   60   46  76.67  Attrib.risk.exp  0.07
## 4                NA <NA>      NA Attrib.risk.pop  0.04
## 5                tira = 0   161 <NA>      NA Risk difference  0.21
## 6                Exposed    17   4  23.53      Risk Ratio  11.29
## 7                Unexposed  144   3   2.08  Attrib.risk.exp  0.91
## 8                NA <NA>      NA Attrib.risk.pop  0.52
## 9                Missing / Missing %   18  6.2%      NA          <NA>      NA
##   95%CI-l1 95%CI-ul
## 1    -0.09   0.21
## 2     0.89   1.30
## 3    -0.12   0.23
## 4      NA     NA
## 5     0.01   0.42
## 6     2.76  46.26
## 7     0.64   0.98
## 8      NA     NA
## 9     NA     NA
##
## $df2
##                Point Estimate  Chi2 p.value  Stats 95%CI-l1 95%CI-ul
## 1      Woolf test of homogeneity 10.47  0.001    NA     NA     NA
## 2                Crude RR for wmouse    NA     NA   2.84   2.12   3.80
## 3      MH RR wmouse adjusted for tira    NA     NA   1.23   1.02   1.48
## 4      Adjusted/crude relative change    NA     NA -56.70   NA     NA
```

**Example 2 : CSInter ill - beer by tira (formatted)**

The following results tables are outputs in “markdown” using the kable function.

```
res <- CSInter(DF, "ill", "beer", "tira", full = TRUE)
```

CSInter ill - beer by(tira)	Total	Cases	Risk %	P.est.	Stats	95%CI-ll	95%CI-ul
tira = 1	116		NA	Risk difference	-0.18	-0.35	-0.01
Exposed	41	27	65.85	Risk ratio	0.78	0.62	1.00
Unexposed	75	63	84.00	Prev. frac. ex.	0.22	0.00	0.38
			NA	Prev. frac. pop	0.08	NA	NA
tira = 0	150		NA	Risk difference	0.00	-0.07	0.07
Exposed	63	3	4.76	Risk Ratio	1.04	0.24	4.47
Unexposed	87	4	4.60	Attrib.risk.exp	0.03	-3.16	0.78
			NA	Attrib.risk.pop	0.01	NA	NA
Missing / Missing %	25	8.6%	NA		NA	NA	NA

	Point Estimate	Chi2	p.value	Stats	95%CI-ll	95%CI-ul
Woolf test of homogeneity		0.14	0.713	NA	NA	NA
Crude RR for beer		NA		0.70	0.49	0.99
MH RR beer adjusted for tira		NA		0.80	0.62	1.03
Adjusted/crude relative change		NA		14.93	NA	NA

### Example 3: CSInter ill - beer by tportion (formatted)

The following results tables are outputs in “markdown” using the kable function.

```
res <- CSInter(DF, "ill", "beer", "tportion", full = TRUE)
kable(res$df1, align="r")
```

CSInter ill - beer by(tportion)	Total	Cases	Risk %	P.est.	Stats	95%CI-ll	95%CI-ul
tportion = 2	53		NA	Risk difference	0.01	-0.16	0.19
Exposed	19	17	89.47	Risk Ratio	1.01	0.83	1.23
Unexposed	34	30	88.24	Attrib.risk.exp	0.01	-0.20	0.19
			NA	Attrib.risk.pop	0.01	NA	NA
tportion = 1	63		NA	Risk difference	-0.35	-0.59	-0.11
Exposed	22	10	45.45	Risk ratio	0.56	0.35	0.91
Unexposed	41	33	80.49	Prev. frac. ex.	0.44	0.09	0.65
			NA	Prev. frac. pop	0.15	NA	NA
tportion = 0	150		NA	Risk difference	0.00	-0.07	0.07
Exposed	63	3	4.76	Risk Ratio	1.04	0.24	4.47
Unexposed	87	4	4.60	Attrib.risk.exp	0.03	-3.16	0.78
			NA	Attrib.risk.pop	0.01	NA	NA
Missing / Missing %	25	8.6%	NA		NA	NA	NA

```
kable(res$df2, align="r")
```

	Point Estimate	Chi2	p.value	Stats	95%CI-ll	95%CI-ul
Woolf test of homogeneity		4.87	0.087	NA	NA	NA
Crude RR for beer		NA		0.70	0.49	0.99
MH RR beer adjusted for tportion		NA		0.80	0.62	1.02
Adjusted/crude relative change		NA		14.62	NA	NA

By storing the results in the object “res”, you are able to use the result table in Markdown as shown above. You can also use individual elements of the results. For example if you would like to view just the Mantel-Haenszel risk ratio for beer adjusted for tportion, you can view it by typing:

```
res$df2$Stats[3]
```

```
## [1] 0.80
## Levels: NA 0.70 0.80 14.62
```

## CC

CC is used for case control studies to determine the association between an exposure and an outcome. Variables need to be binary and coded as “0” and “1”. Point estimates and confidence intervals for the odds ratio are calculated along with attributable or preventive fractions for the exposed and total population. Additionally you can select if you want to display the Fisher’s exact test, by specifying `exact = TRUE`. If you specify `full = TRUE` you can easily access useful statistics from the output tables.

### Syntax

```
CC(x, cases, exposure, exact, full=FALSE)
```

### Example 1: CC ill - mousse (unformatted)

```
cc(DF, "ill", "mousse", exact = TRUE)
```

```
## $df1
##           Cases Controls Total
## Exposed           81      42   123
## Unexposed          22     144   166
## Total             103     186   289
## Proportion exposed 0.79     0.23 0.43
##
## $df2
##           Point estimate 95%CI-l1 95%CI-ul
## Odds ratio           12.62     6.80   23.70
## Attr. frac. ex.         0.92     0.85   0.96
## Attr. frac. pop         0.72     NA     NA
## chi2(1)                85.22     NA     NA
## Pr>chi2                 0.000     NA     NA
## Fisher p-value         0.000     NA     NA
```



## Example 2: CC ill - beer (formatted)

The following results tables are outputs in “markdown” using the kable function.

```
result <- CC(DF, "ill", "beer", exact = TRUE, full = TRUE)
kable(result$df1, align="r")
```

	Cases	Controls	Total
Exposed	30	76	106
Unexposed	69	96	165
Total	99	172	271
Proportion exposed	0.30	0.44	0.39

```
kable(result$df2, align=result$df2.align)
```

	Point estimate	95%CI-ll	95%CI-ul
Odds ratio	0.55	0.31	0.95
Prev. frac. ex.	0.45	0.05	0.69
Prev. frac. pop	0.20	NA	NA
chi2(1)	5.09	NA	NA
Pr>chi2	0.024	NA	NA
Fisher p-value	0.028	NA	NA

By storing the results in the object “result”, you are able to use the result tables in Markdown as shown above. By specifying “full = TRUE” you can also easily use individual elements of the results. For example if you would like to view just the odds ratio, you can view it by typing:

```
result$st$odds_ratio$point_estimate
```

```
## [1] 0.5491991 0.3127957 0.9547369
```

## CCTable - Summary table for case control studies

CCTable is used for univariate analysis of case control studies with several exposures. The results are summarised in one table with one row per exposure making comparisons between exposures easier and providing a useful table for integrating into reports. Note that all variables need to be numeric and binary and coded as “0” and “1”.

The results of this function contain: The name of exposure variables, the total number of cases, the number of exposed cases, the percentage of exposed among cases, the number of controls, the number of exposed controls, the percentage of exposed among controls, odds ratios, 95%CI intervals, p-values.

You can optionally choose to display the Fisher’s exact p-value instead of the Chi squared p-value, with the option `exact = TRUE`.

You can specify the sort order, with the option `sort=“or”` to order by odds ratios. The default sort order is by p-values.

The option `“full = TRUE”` provides you with useful formatting information, which can be handy if you’re using “markdown”.

### Syntax

```
CCTable(x, cases, exposure=c(), exact=FALSE, sort = “pvalue”, full=FALSE)
```

### Example 1: CCTable results ordered by p-value (unformatted)

```
CCTable(DF, "ill",
  exposure = c("sex", "agegroup", "tira", "beer", "mousse", "wmousse", "dmousse",
    "redjelly", "fruitsalad", "tomato", "mince", "salmon", "horseradish",
    "chickenwin", "roastbeef", "pork"))
```

```
## $df
##
## Tot.Cases Exposed % Tot.Ctrls Exposed % OR CI ll
## tira 101 94 93.07 185 27 14.59 78.58 31.45
## mousse 103 81 78.64 186 42 22.58 12.62 6.80
## wmousse 98 49 50.00 179 23 12.85 6.78 3.62
## dmousse 102 76 74.51 185 37 20.00 11.69 6.36
## redjelly 103 45 43.69 188 34 18.09 3.51 1.98
## fruitsalad 103 46 44.66 188 25 13.30 5.26 2.86
## beer 99 30 30.30 172 76 44.19 0.55 0.31
## tomato 103 35 33.98 188 48 25.53 1.50 0.86
## pork 102 48 47.06 187 72 38.50 1.42 0.85
## horseradish 102 30 29.41 187 42 22.46 1.44 0.80
## sex 103 50 48.54 188 102 54.26 0.80 0.48
## roastbeef 103 8 7.77 188 21 11.17 0.67 0.25
## chickenwin 103 33 32.04 188 51 27.13 1.27 0.72
## mince 103 32 31.07 188 55 29.26 1.09 0.62
## agegroup 100 25 25.00 183 43 23.50 1.09 0.59
## salmon 100 37 37.00 187 67 35.83 1.05 0.61
##
## CI ul p(Chi2)
## tira 217.15 0.000
## mousse 23.70 0.000
## wmousse 12.83 0.000
## dmousse 21.64 0.000
## redjelly 6.24 0.000
## fruitsalad 9.75 0.000
## beer 0.95 0.024
## tomato 2.61 0.127
## pork 2.38 0.158
## horseradish 2.57 0.192
## sex 1.32 0.351
## roastbeef 1.65 0.354
## chickenwin 2.20 0.377
```

## mince	1.89	0.747
## agegroup	1.98	0.777
## salmon	1.79	0.844

## Example 2: CCTable results ordered by odds ratio (formatted)

The following results tables are outputs in “markdown” using the kable function.

```
res = CCTable(DF, "ill", sort = "or", exposure = Colnames)
kable(res$df)
```

	Tot.Cases	Exposed	%	Tot.Ctrls	Exposed	%	OR	CI ll	CI ul	p(Chi2)
tira	101	94	93.07	185	27	14.59	78.58	31.45	217.15	0.000
mousse	103	81	78.64	186	42	22.58	12.62	6.80	23.70	0.000
dmousse	102	76	74.51	185	37	20.00	11.69	6.36	21.64	0.000
wmousse	98	49	50.00	179	23	12.85	6.78	3.62	12.83	0.000
fruitsalad	103	46	44.66	188	25	13.30	5.26	2.86	9.75	0.000
redjelly	103	45	43.69	188	34	18.09	3.51	1.98	6.24	0.000
tomato	103	35	33.98	188	48	25.53	1.50	0.86	2.61	0.127
horseradish	102	30	29.41	187	42	22.46	1.44	0.80	2.57	0.192
pork	102	48	47.06	187	72	38.50	1.42	0.85	2.38	0.158
chickenwin	103	33	32.04	188	51	27.13	1.27	0.72	2.20	0.377
mince	103	32	31.07	188	55	29.26	1.09	0.62	1.89	0.747
agegroup	100	25	25.00	183	43	23.50	1.09	0.59	1.98	0.777
salmon	100	37	37.00	187	67	35.83	1.05	0.61	1.79	0.844
sex	103	50	48.54	188	102	54.26	0.80	0.48	1.32	0.351
roastbeef	103	8	7.77	188	21	11.17	0.67	0.25	1.65	0.354
beer	99	30	30.30	172	76	44.19	0.55	0.31	0.95	0.024

### Example 3: CCTable results ordered by p-value from the Fisher's exact test (formatted)

The following results tables are outputs in "markdown" using the kable function.

```
res = CCTable(DF, "ill", exposure = Colnames, exact=TRUE)
kable(res$df)
```

	Tot.Cases	Exposed	%	Tot.Ctrls	Exposed	%	OR	CI ll	CI ul	p(Fisher)
tira	101	94	93.07	185	27	14.59	78.58	31.45	217.15	0.000
wmousse	98	49	50.00	179	23	12.85	6.78	3.62	12.83	0.000
dmousse	102	76	74.51	185	37	20.00	11.69	6.36	21.64	0.000
mousse	103	81	78.64	186	42	22.58	12.62	6.80	23.70	0.000
redjelly	103	45	43.69	188	34	18.09	3.51	1.98	6.24	0.000
fruitsalad	103	46	44.66	188	25	13.30	5.26	2.86	9.75	0.000
beer	99	30	30.30	172	76	44.19	0.55	0.31	0.95	0.028
tomato	103	35	33.98	188	48	25.53	1.50	0.86	2.61	0.137
pork	102	48	47.06	187	72	38.50	1.42	0.85	2.38	0.171
horseradish	102	30	29.41	187	42	22.46	1.44	0.80	2.57	0.203
sex	103	50	48.54	188	102	54.26	0.80	0.48	1.32	0.391
roastbeef	103	8	7.77	188	21	11.17	0.67	0.25	1.65	0.417
chickenwin	103	33	32.04	188	51	27.13	1.27	0.72	2.20	0.418
agegroup	100	25	25.00	183	43	23.50	1.09	0.59	1.98	0.773
mince	103	32	31.07	188	55	29.26	1.09	0.62	1.89	0.789
salmon	100	37	37.00	187	67	35.83	1.05	0.61	1.79	0.898

By storing the results in the object "res", you are able to use the result table in Markdown as shown above. You can also use individual elements of the results. For example if you would like to view just the odds ratio, you can view it by typing (for example):

```
res$df$OR[1]
```

```
## [1] 78.58
```

```
## 15 Levels: 0.55 0.67 0.80 1.05 1.09 1.27 1.42 1.44 1.50 11.69 ... 78.58
```

## CCInter - Stratified analysis for case control studies

CCInter is useful to determine the effects of a third variable on the association between an exposure and an outcome. CCInter produces 2 by 2 tables with stratum specific odds ratios, attributable risk among exposed and population attributable risk.

Note that the outcome and exposure variable need to be numeric and binary and coded as "0" and 1". The third variable needs to be numeric, but may have more categories, such as "0", "1" and "2".

CCInter displays a summary with the crude OR, the Mantel Haenszel adjusted OR and the result of a Woolf test for homogeneity of stratum-specific OR.

The option "full = TRUE" provides you with useful formatting information, which can be handy if you're using "markdown".

### Syntax

```
CCInter (x, cases, exposure, by, full=FALSE)
```

### Example 1: CCInter ill - wmousse by tira (unformatted)

```
CCInter(DF, cases="ill", exposure = "wmousse", by = "tira")

## $df1
##   CCInter ill - wmousse by(tira) Cases Controls      P.est. Stats
## 1                tira = 1 <NA>    <NA>      Odds ratio 1.45
## 2                Exposed   43      9  Attrib.risk.exp 0.31
## 3                Unexposed  46     14  Attrib.risk.pop 0.15
## 4                Total    89     23
## 5                Exposed % 48.3% 39.1%
## 6                <NA>      <NA>
## 7                tira = 0 <NA>    <NA>      Odds ratio 14.46
## 8                Exposed   4     13  Attrib.risk.exp 0.93
## 9                Unexposed  3    141  Attrib.risk.pop 0.53
## 10               Total    7    154
## 11               Exposed % 57.1%  8.4%
## 12               <NA>      <NA>
## 13               Number of obs 273 <NA>
## 14               Missing    18 <NA>
## 95%CI-l1 95%CI-ul
## 1      0.52    4.22
## 2     -0.92    0.76
## 3      <NA>   <NA>
## 4      <NA>   <NA>
## 5      <NA>   <NA>
## 6      <NA>   <NA>
## 7      2.12  106.00
## 8      0.53    0.99
## 9      <NA>   <NA>
## 10     <NA>   <NA>
## 11     <NA>   <NA>
## 12     <NA>   <NA>
## 13     <NA>   <NA>
## 14     <NA>   <NA>
##
## $df2
##                P.estimate  Stats 95%CI-l1 95%CI-ul
## 1 MH test of Homogeneity (p-value) 0.01
## 2           Crude OR for wmousse  6.76    3.57    12.93
## 3 MH OR wmousse adjusted for tira  2.25    1.01    5.05
## 4 Adjusted/crude relative change -66.65    -    -
```

## Example 2: CCInter ill - beer by tira (formatted)

The following results tables are outputs in “markdown” using the kable function.

```
res <- CCInter(DF, cases="ill", exposure = "beer", by = "tira", full = TRUE)
kable(res$df1, align=res$df1.align)
```

CCInter ill - beer by(tira)	Cases	Controls	P.est.	Stats	95%CI-ll	95%CI-ul
tira = 1			Odds ratio	0.37	0.14	0.99
Exposed	27	14	Prev. frac. ex.	0.63	0.01	0.86
Unexposed	63	12	Prev. frac. pop	0.34		
Total	90	26				
Exposed %	30.0%	53.8%				
<hr/>						
tira = 0			Odds ratio	1.04	0.15	6.38
Exposed	3	60	Attrib.risk.exp	0.04	-5.82	0.84
Unexposed	4	83	Attrib.risk.pop	0.02		
Total	7	143				
Exposed %	42.9%	42.0%				
<hr/>						
Number of obs	266					
Missing	25					

```
kable(res$df2)
```

P.estimate	Stats	95%CI-ll	95%CI-ul
MH test of Homogeneity (p-value)	0.22		
Crude OR for beer	0.57	0.33	1.00
MH OR beer adjusted for tira	0.48	0.22	1.05
Adjusted/crude relative change	-15.83	—	—

### Example 3: CCIInter ill - beer by tportion (formatted)

The following results tables are outputs in “markdown” using the kable function.

```
res <- CCIInter(DF, cases="ill", exposure = "beer", by = "tportion", full = TRUE)
kable(res$df1, align=res$df1.align)
```

CCIInter ill - beer by(tportion)	Cases	Controls	P.est.	Stats	95%CI-l	95%CI-ul
tportion = 2			Odds ratio	1.13	0.14	13.73
Exposed	17	2	Attrib.risk.exp	0.12	-5.94	0.93
Unexposed	30	4	Attrib.risk.pop	0.04		
Total	47	6				
Exposed %	36.2%	33.3%				
<hr/>						
tportion = 1			Odds ratio	0.20	0.06	0.73
Exposed	10	12	Prev. frac. ex.	0.80	0.27	0.94
Unexposed	33	8	Prev. frac. pop	0.48		
Total	43	20				
Exposed %	23.3%	60.0%				
<hr/>						
tportion = 0			Odds ratio	1.04	0.15	6.38
Exposed	3	60	Attrib.risk.exp	0.04	-5.82	0.84
Unexposed	4	83	Attrib.risk.pop	0.02		
Total	7	143				
Exposed %	42.9%	42.0%				
<hr/>						
Number of obs	266					
Missing	25					

```
kable(res$df2, align=res$df2.align)
```

P.estimate	Stats	95%CI-l	95%CI-ul
MH test of Homogeneity (p-value)	0.13		
Crude OR for beer	0.57	0.33	1.00
MH OR beer adjusted for tportion	0.47	0.21	1.02
Adjusted/crude relative change	-18.73	—	—

By storing the results in the object “res”, you are able to use the result table in Markdown as shown above. You can also use individual elements of the results. For example if you would like to view just the Mantel-Haenszel odds ratio for beer adjusted for tportion, you can view it by typing:

```
##res$df2$Stats[3]
```

```
)
```