Microsoft Excel add-in for the statistical analysis of contingency tables

Peter Slezák
Institute of Simulation and Virtual Medical Education, Faculty of Medicine, Comenius University, Špitálska 24, 813 72 Bratislava, Slovakia
peter.slezak@fmed.uniba.sk

Pavol Bokes
Department of Applied Mathematics and Statistics, Faculty of Mathematics, Physics and Informatics, Comenius University, Mlynská dolina, 842 48 Bratislava, Slovakia
bokes@pc2.iam.fmph.uniba.sk

Pavol Námer
Polymer Institute, Slovak Academy of Sciences, Dúbravská cesta 9, 845 41 Bratislava, Slovakia
pavol.namer@savba.sk

Iveta Waczulíková
Division of Biomedical Physics, Faculty of Mathematics, Physics and Informatics, Comenius University, Mlynská dolina, 842 48 Bratislava, Slovakia
waczulikova@fmph.uniba.sk

Abstract

This paper introduces "Contingency table analysis", a freely available menu-driven add-in program for Microsoft EXCEL, written in Visual Basic for Applications (VBA), for basic univariate and bivariate statistical analyses of contingency tables. The program provides modules for the statistical analysis of proportions, 2 × 2 tables, stratified 2 × 2 tables, and R × C tables. We compare the results of the analyses performed using our software with those obtained by commercially available statistical software. The comparison shows that our software performs equally well. The use of the add-in facilitates the convenient prosecution of basic statistical analyses on contingency tables from within EXCEL, sparing us the additional cost, or the inconvenience of alternating between multiple platforms, often incurred in using a commercial statistical package.

1. Introduction

Microsoft EXCEL is a versatile spreadsheet application that is widely used by clinicians, biomedical scientists and students. MS EXCEL offers a wide range of applications, ranging from clinical data management and simple analysis [1,2], to collecting data from clinics to a central database [3,4]. In biomedical research there are two broad categories of studies that produce statistical data: the first one is designed controlled experiments and the second one is observational studies. Researchers often record events, counts and/or proportions to provide evidence for associations between characteristics of certain populations, e.g. between a possible risk factor and a disease. MS EXCEL provides tools for simple statistical analysis, such as t-tests, F-test, ANOVA, correlation and ordinary least squares (OLS) regression. Moreover, EXCEL's built-in analytical tools have been extended by add-ins useful for pharmacokinetics analysis [5,6], microarray data (statistical) analysis [7,8] , multiple comparison procedures [9] and data presentation [10].

MS EXCEL also allows for a convenient representation of categorical data and creating cross-classification tables (two-way/two-dimensional contingency tables) from raw data. Nevertheless, the software lacks methods required for the analysis of these tables. The only exception is the CHISQ.TEST function for determining whether the observed data differ from the expected. This test is, however, cumbersome to perform since it requires previous calculation of the expected frequencies. Since the use of categorical data testing procedures is necessary for assessing the significance of association between the characteristics (e.g. disease and a risk factor in a 2×2 table), the researchers need to use additional statistical software.

Our aim was to create an add-in that would fill the gap and perform basic univariate and bivariate statistical analyses of contingency tables from within EXCEL in a comfortable way, sparing us the additional cost or the inconvenience of alternating between multiple platforms, often incurred in using a commercial statistical package. The developed tools are described in more detail in the text below.

2. Computational methods and theory

2.1. Two-way contingency tables

Data on two discrete variables, say *X* and *Y*, are commonly recorded in a contingency table. The table has r rows and c columns, where r is the number of (all possible) categories for *X* and c is the number of categories for *Y*. The cell in the *i*-th row and *j*-th column gives the number n_{ij} of experimental units which are classified into the *i*-th category in *X* and *j*-th category in *Y*. The $r \times c$ entries of the table are typically flanked by the marginal frequencies,

$$n_{i,\cdot} = \sum_{j=1}^{c} n_{i,j}, \quad n_{\cdot,i} = \sum_{i=1}^{r} n_{i,j}$$

and by the labels for the row and column categories (cf. Table 1). The sample size

$$n_{\cdot,\cdot} = \sum_{i=1}^{r} \sum_{j=1}^{c} n_{i,j}$$

is entered in the bottom-right corner of the table.

The data recorded in contingency tables can be collected in two ways (at least). First, a random sample can be drawn from a joint population distribution of *X* and *Y*; such design implies that both row and column marginal frequencies are random. Alternatively, the population of interest is stratified by the variable *X*, and a random sample is drawn from the distribution of *Y* in each stratum; in this case the marginal frequencies of *X* are fixed by the experimenter [11].

Table 1. The layout of a r × c contingency table.

i	*j*				Σ
	1	*2*	...	*c*	
1	$n_{1,1}$	$n_{1,2}$	...	$n_{1,c}$	$\mathbf{n_{1,\cdot}}$
2	$n_{2,1}$	$n_{2,2}$	...	$n_{2,c}$	$\mathbf{n_{2,\cdot}}$
⋮	⋮	⋮	⋮	⋮	⋮
r	$n_{r,1}$	$n_{r,2}$	...	$n_{r,c}$	$\mathbf{n_{r,\cdot}}$
Σ	$\mathbf{n_{\cdot,1}}$	$\mathbf{n_{\cdot,2}}$	...	$\mathbf{n_{\cdot,c}}$	$\mathbf{n_{\cdot,\cdot}}$

In the former case, the null hypothesis for the lack of association between *X* and *Y* is formulated by the symmetric relation

$$H_0: \pi_{i,j} = \pi_{i,\cdot}\pi_{\cdot,j}\ , \qquad (1)$$

requiring that the joint population distribution of *X* and *Y* be a product of its marginals. In the latter situation, the null hypothesis is conveniently formulated as

$$H_0: \pi_{j|i} = \pi_{j|i'}, \text{ for any pair i and i'}, \qquad (2)$$

requiring that the conditional distribution of *Y* be the same in each stratum. In either case, we can conclude that there is no evidence for an association between the two variables if the null hypothesis cannot be rejected.
The test statistic for the hypotheses (1) or (2) is defined by

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(n_{i,j} - \hat{n}_{i,j}\right)^2}{\hat{n}_{i,j}}, \qquad (3)$$

in which

$$\hat{n}_{i,j} = \frac{n_{i,\cdot}n_{\cdot,j}}{n_{\cdot,\cdot}}$$

are the expected frequencies if the null hypothesis is true.
In the limit of large sample sizes, the asymptotic distribution of the test statistic is χ^2 with $(r - 1)(c - 1)$ degrees of freedom [12]. The null hypothesis is rejected for large values of the statistic.
According to Cochran's rule, the χ^2 test can be used provided that all expected frequencies are greater than one, and fewer than a fifth of them are less than five [12]. If these criteria are not met, the use of the Fisher exact test is recommended in 2 × 2 tables [12].

2.2. Measures of association

If we reject the null hypotheses (1) or (2), concluding that there is a statistically significant association between the row and column variables, we may wish to characterise the degree of association by a single numerical quantity. Among the various coefficients of association available, some are defined as a reparametrization of the χ^2 statistic, adjusting for the sample size (Pearson's contingency coefficient *C*, Cramer's *V* and coefficient Φ); another class of coefficients is appropriate if the categories for the two variables are ranked (Goodman-Kruskal's γ, Kendall's τ_b or τ_c). All of the above coefficients are less than one in absolute value, with the zero value corresponding to a lack of association [12].

2.3. Double dichotomy

In clinical practice, we are often interested in testing for association between the exposure to a risk factor (modelled by a dichotomous variable *X*) and the incidence of a disease (a dichotomous variable *Y*) in a population of individuals. Data on double dichotomy are reported in a 2 × 2 table.

The null hypothesis of no association is tested by the χ^2 statistic with one degree of freedom or by the Fisher exact test, should the expected frequencies be low, as described previously.

There are several quantities by which the level of association between two dichotomous variables can be characterised.

- The difference

$$\mathrm{D} = \mathrm{P}[\text{sick}|\text{exposed}] - \mathrm{P}[\text{sick}|\text{not exposed}]$$

can be interpreted as the percentage of disease incidents that can be attributed to the risk factor.

- The risk ratio

$$\mathrm{RR} = \frac{\mathrm{P}[\text{sick}|\text{exposed}]}{\mathrm{P}[\text{sick}|\text{not exposed}]}$$

indicates how many times more (or less) likely the disease occurs in the exposed population than in the unexposed.

- The odds ratio

$$\mathrm{OR} = \frac{\mathrm{P}[\text{sick}|\text{exposed}]/\mathrm{P}[\text{healthy}|\text{exposed}]}{\mathrm{P}[\text{sick}|\text{not exposed}]/\mathrm{P}[\text{healthy}|\text{not exposed}]}$$

overcomes an asymmetry in the definition of the risk ratio, which is sensitive to an increase in incidence of rare diseases, but is insensitive to a decrease in avoidance of common diseases; the odds ratio is sensitive to both, being approximately equal to the risk ratio for rare diseases.

The above quantities can be estimated from the frequency table as

$$\widehat{\mathrm{D}} = \frac{n_{1,1}}{n_{1,\cdot}} - \frac{n_{2,1}}{n_{2,\cdot}}, \quad \widehat{\mathrm{RR}} = \frac{n_{1,1}n_{2,\cdot}}{n_{2,1}n_{1,\cdot}}, \quad \widehat{\mathrm{OR}} = \frac{n_{1,1}n_{2,1}}{n_{1,2}n_{2,2}},$$

in which 1 stands for the presence of the disease (or exposure) and 2 for their absence.

2.4. Testing for linear trend in 2 × c tables

If the row variable X is dichotomous, and the column variable Y is ranked, we may wish to inquire whether there is a linear trend, as the column index j increases, in the proportions $n_{1,j}/n_{\cdot,j}$ of individuals in the first row. The column categories are allotted an increasing, typically equally spaced, sequence of scores $x_1 < x_2 \ldots < x_c$. The probabilities $\pi_{1|j}$ are estimated by a simple linear regression of the proportions $n_{1,j}/n_{\cdot,j}$ on x_i, weighted by the column totals $n_{\cdot,j}$ [14].

Comparing the frequencies thus estimated to the frequencies expected if the null hypothesis of no association was true, one obtains a goodness-of-fit test statistic, such as in (3), which here reduces to

$$\chi^2 = \frac{n_{\cdot,\cdot}\left(n_{\cdot,\cdot}\sum_{i=1}^{c} n_{i,1}\,x_i - n_{\cdot,1}\sum_{i=1}^{c} n_{i,\cdot}x_i\right)^2}{n_{\cdot,1}\left(n_{\cdot,\cdot} - n_{\cdot,1}\right)\left[n_{\cdot,\cdot}\sum_{i=1}^{c} n_{i,\cdot}x_i^2 - \left(\sum_{i=1}^{c} n_{i,\cdot}x_i\right)^2\right]}. \qquad (4)$$

If the null hypothesis is true, the asymptotic distribution of the test statistic is χ^2 with one degree of freedom. The hypothesis is rejected for large values of the statistic.

In the test for trend in proportions, as opposed to the general χ^2 test for association, the restricted nature of the alternative - that the proportions $\pi_{1|j}$ depend linearly on the scores x_i - leads to a reduction, by c - 2, in the degrees of freedom of the test statistic. Consequently, the test for trend can be more powerful that the general test for association, revealing trends that may have otherwise passed unnoticed [12].

2.5. Stratified 2 × 2 tables (2 × 2 × k tables)

In 2 × 2 tables, the row variable (exposure to a risk factor) and the column variable (incidence of a disease) can both be associated with a third variable Z (e.g. another risk factor). The association between X and Y can be either exaggerated or undervalued by the confounding factor Z (Simpson's paradox) [15]. The Mantel-Haenszel

test for association between *X* and *Y* adjusts for the effect of the confounding variable. The test statistic is given by

$$\chi^2 = \frac{\left(\left|\sum_{i=1}^{k} n_{1,1}(i) - \sum_{i=1}^{k} \frac{n_{1,\cdot}(i)n_{\cdot,1}(i)}{n_{\cdot,\cdot}(i)}\right| - \frac{1}{2}\right)^2}{\sum_{i=1}^{k} \frac{n_{1,\cdot}(i)n_{2,\cdot}(i)n_{\cdot,1}(i)n_{\cdot,2}(i)}{n_{\cdot,\cdot}(i)^3 - n_{\cdot,\cdot}(i)^2}}, \quad (5)$$

If the null hypothesis of no association is true, the asymptotic distribution of the test statistic is χ^2 with one degree of freedom [16].

The overall strength of the association can be characterised by a pooled odds ratio, which is estimated by

$$\widehat{OR}_{MH} = \sum_{i=1}^{k} \frac{n_{1,1}(i)n_{2,2}(i)}{n_{\cdot,\cdot}(i)} \Big/ \sum_{i=1}^{k} \frac{n_{1,2}(i)n_{2,1}(i)}{n_{\cdot,\cdot}(i)}, \quad (6)$$

The confidence interval for the pooled odds ratio can be obtained using an estimate of the variance of the logarithm of the above estimate (see e.g. [16]).

Table 2. The modules with the summary description of the implemented methods. RR – risk ratio

Module	Implemented methods	Computed outcomes
2x2 table	Fisher exact test	One- and two sided P-values and mid-P values
	χ^2 test of independence	Test statistic (No continuity correction), two sided P-value
	χ^2 measures of nominal association	Cramer's *V*, Coefficient *Φ*, Contingency coefficient
	Odds ratio, Risk ratio	Point estimate with 95% CI
	Paired data	Liddell's test, RR with 95% CI
RxC table	χ^2 test of independence	Test statistic, two sided P-value, Monte Carlo estimate of the exact P-value with 99% CI
	χ^2 measures of nominal association	Cramer's *V*, Coefficient *Φ*, Contingency coefficient
	measures of ordinal association	Kendall's π_b and π_c, Goodman-Kruskal γ Point estimates of these coefficients with 95%CI
	Cochran-Armitage test	Test for linear trend in 2 × C table; test for departure from linear trend
Mantel-Haenszel	Mantel-Haenszel test	Mantel-Haenszel test; Pooled OR with 95% CI
Proportions	Single	Point estimate with 95% CI
	Two independent	Point estimate with 95% CI of the difference between proportions
	Paired proportions	

3. Program Description

The The EXCEL add-in is written entirely in EXCEL Visual Basic for Applications (VBA). The *Contingency table analysis* add-in can be installed automatically by double clicking on the .xlam file or manually like other EXCEL add-ins. Macros need to be enabled in EXCEL to successfully install and run this Add-in. The list of the implemented methods is summarised in Table 2.

Once the add-in program has been installed, a pull-down <Contingency table analysis > menu appears in the Add-in tab after EXCEL is launched. As shown in Fig. 1, users can select a module of interest from the menu. Input data can be specified by simply drag-selecting the range of cells in the spreadsheet (see Fig. 2,3,4), alternatively the modules <2×2 table> and <Proportions> enable manual input of data. The calculation options can be set interactively by clicking on the requested option in the dialogue window. Each module automatically creates a short tabular report with results either in a new workbook, worksheet or in a specified range of cells.

In addition, we created a help file which provides the user with contextual information about data input, implemented methods, and further references.

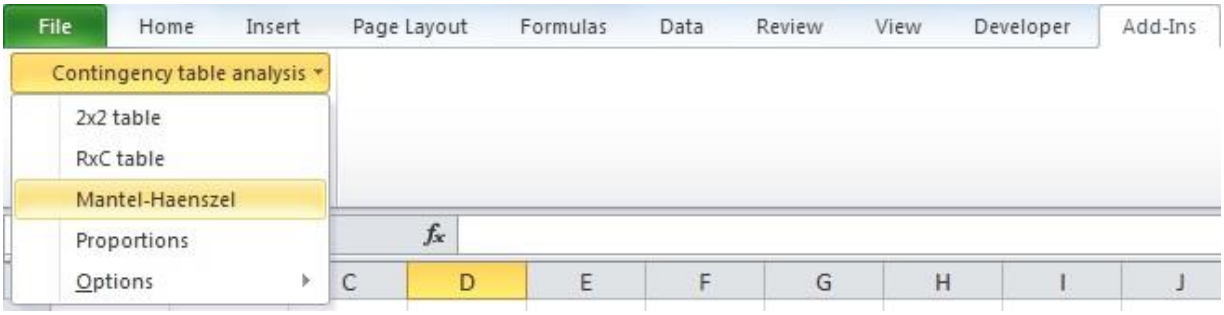

Fig. 1 - Contingency table analysis menu in MS EXCEL Add-Ins tab.

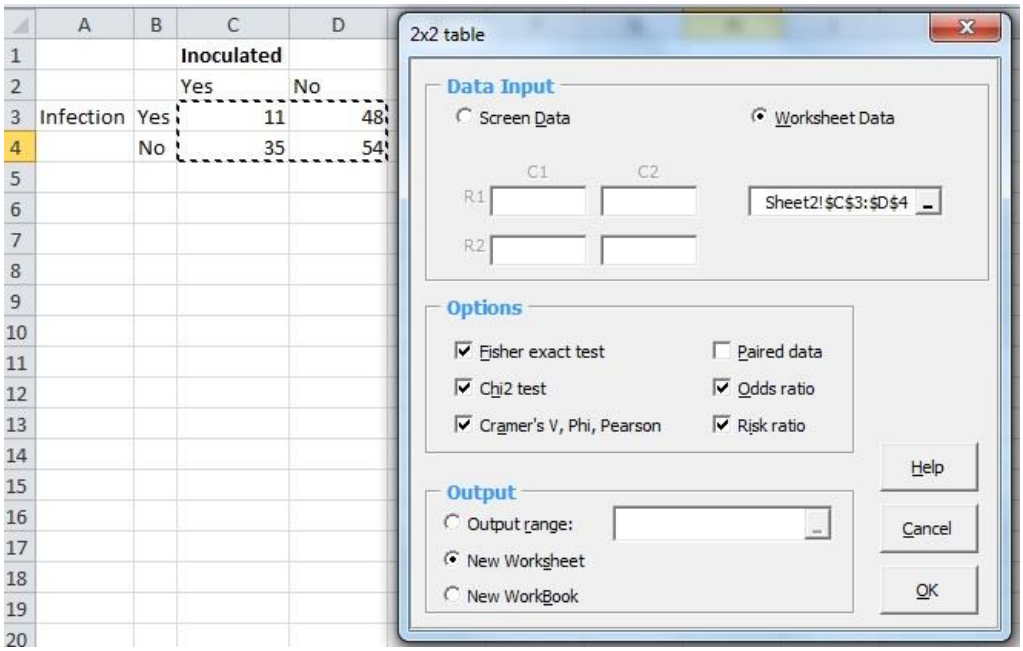

Fig. 2 - 2 × 2 table module dialogue window.

4. Sample of program runs

We recommend using a Pivot Table tool for contingency table preparations from large data sets. This add-in was created to equip EXCEL with a tool for statistical analysis of contingency tables. It is suitable for analysis of small-to-medium-size data tables.

Example data sets used in this section can be found at http://bio-med-stat.webnode.sk/ms-excel-add-ins/contingency-table-analysis-addin/

4.1. 2 × 2 table analysis

Consider a study investigating the incidence of a virus infection (outcome) recorded in two groups, one group having been previously inoculated (exposed) and the other group without inoculation (unexposed) [17]. The input of the data and available computing options are presented in Fig 2.

Table 3 shows outputs of the analyses performed on the table presented in Fig. 2 with the *Contingency table analysis* add-in, with StatsDirect 2.7.9 (StatsDirect Ltd. StatsDirect statistical software http://www.statsdirect.com) and GraphPad Prism 6 for Windows (GraphPad Software, San Diego, California, USA, www.graphpad.com). Results are rounded to six decimal places. We can see a complete agreement in the

results obtained by the three programs for all calculated statistics except for OR and RR 95% confidence intervals: however, the differences are negligible. Another way of looking at these data is to consider them as two independent proportions. Therefore, we also use the Proportions module to calculate the difference in proportions of infected individuals in the inoculated and not inoculated groups (Tab. 3). Results are comparable with those provided by Statsdirect, which, however, uses a different method for the estimation of confidence intervals.

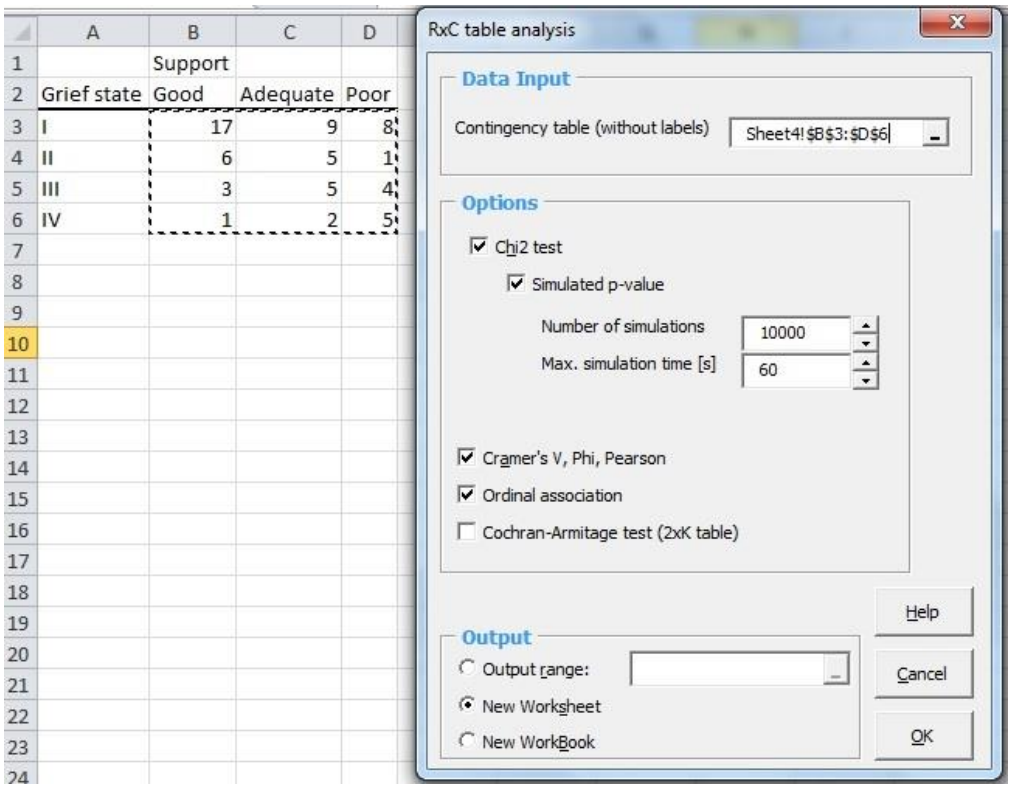

Fig. 3 - R × C table module dialogue window.

Table 3 - Comparison of outputs of the analyses provided by *Contingency table addin* – 2 × 2 table module, and statistical software Statsdirect 2.7.9 and GraphPad Prism 6.

	Contingency table Add-in	Statsdirect	GraphPad Prism
One sided P-value[i]	0.005861	0.005861	0.0059
Two sided P-value[i]	0.010712	0.010712	0.0107
mid-P one sided[i]	0.003856	0.003856	n.a.
mid-P two sided[i]	0.007713	0.007713	n.a.
Uncorrected χ^2	7.084698	7.084698	7.084698
P-value (χ^2)	0.007774	0.007774	0.0078
Cramer's *V*	0.218791	-0.218791[ii]	n.a.
Contingency coefficient	0.213735	0.213735	n.a.
Phi	0.218791	0.218791	n.a.
OR (95% CI)[iv]	0.353571 0.161896 to 0.772181	0.353571 0.161898 to 0.77217	0.353571 0.161867 to 0.77232
RR (95% CI)	0.508152 0.291698 to 0.885226	0.508152 0.285906 to 0.849752[iii]	0.508152 0.291661 to 0.885339
Proportions difference (95% CI)	-0.23146 -0.370201 to -0.062625[V]	-0.23146 -0.375537 to -0.062889[iii]	n.a.

[i] values for the Fisher's exact test (GraphPad Prism provides four decimal places only); [ii] Statsdirect computes a signed value; [iii] Miettinen-Nurminen confidence interval; [iv] confidence intervals based on Woolf method; [V] output computed by Proportions module; n.a. - not available.

4.2. R × C table analysis

Consider data collected from 66 mothers, who had suffered a death of their newborn baby [18]. The aim of the study was to assess the relationship between their state of grief and a degree of support offered by the ordinal association option from the R×C table module was used to analyse these data (Fig. 3). The default output is presented in Tab 4. The values of coefficients tau-b, tau-c, gamma and their 95%CIs and P-values agree to at least 5 decimal places with those computed using the commercial software mentioned above (comparison not presented). A Monte Carlo estimation of the Pearson χ^2 statistics based on 10 000 sampled tables leads to P-value of 0.1279 and 99% CI (0.119296 to 0.136504), which is very close to the P-value computed by IBM SPSS 20 (0.1288 and 99% CI from 0.120172 to 0.137428). (It takes only approximately 0.855 seconds to perform this calculation on the 1.4 GHz Intel Centrino 2, Windows 7, 64 bit EXCEL 2010.)

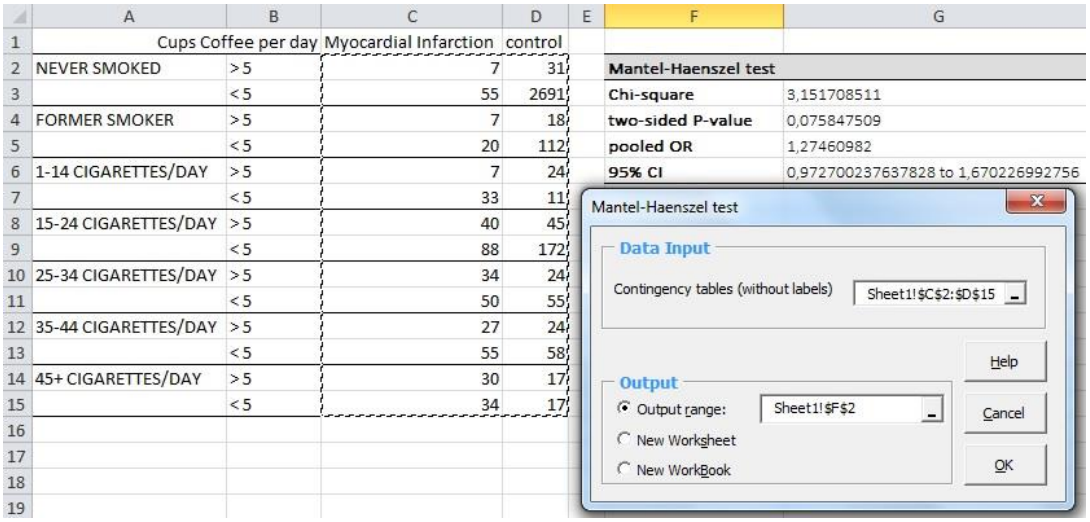

Fig. 4 - Required data arrangement for calculation of the Mantel-Haenszel test.

Table 4 – Default ordinal association output of R × C table module.

Measures of ordinal association	
Kendall's tau-b	0.236078
95% CI	0.019299 to 0.452856
two-sided P-value	0.0328025
Kendall's tau-c	0.232094
95% CI	0.018973 to 0.445214
two-sided P-value	0.032802
Goodman-Kruskal's gamma	0.349223
95% CI	0.028549 to 0.669897
two-sided P-value	0.032802

4.3. Stratified 2 × 2 table analysis

To demonstrate the application of the Mantel-Haenszel module, we evaluate the association between coffee drinking and myocardial infarction in the population of young women aged 30–49 years, where the relationship is suspected to be confounded with smoking (Fig. 4) [19]. The Mantel-Haenszel module provides the pooled odds ratio (OR) and the 95% CI; the respective values are equal to 1.27461 and (0.972700 to 1.670227). The χ^2 statistic for testing whether the OR is different from one amounts to 3.151709 (P = 0.075848). The results agree to at least five decimal places with those computed by Statsdirect that gave the pooled OR of 1.27461 with 95% CI (0.972705 to 1.670219) and χ^2 = 3.151709 (P = 0.0758475).
Based on these three realistic examples, which in our experience are pertinent to clinical practice, we conclude that the *Contingency table analysis* add-in computes the statistics of interest with a highly satisfactory accuracy.

5. Hardware and software specifications

The system and hardware requirements are identical to the requirements for MS EXCEL. MS EXCEL for PC, version 2007 or newer, must be installed on the computer (Office for Mac is not supported). Both 32 and 64 bit

EXCEL versions are supported. The help file has to be located in the same folder as the *Contingency table analysis* add-in file in order to be correctly recognised by the add-in.

6. Availability of the program

The Contingency table analysis add-in file can be obtained either on an e-mail request from the authors, or it can be downloaded from the authors' website http://bio-med-stat.webnode.sk/ms-excel-add-ins/contingency-table-analysis-addin/, or from the Journal pages as a supplementary material to this paper.

7. Acknowledgement

The work was partially supported by the grants KEGA 003UK-4/2012 and VEGA 2/0101/12; PS was partially financially supported by BASF (Slovakia) - The Chemical Company in the form of Preveda award for young scientist; PB gratefully acknowledges the support of the Slovak Research and Development Agency (contract no. APVV-0134-10) and also of the VEGA grant agency (contract no. 1/0711/12). We would also like to thank Oliver Waczulík for assistance in preparing the final text.

10. References

[1] P.P. Gomes, L.A. Passeri, J.R. Barbosa, A 5-year retrospective study of zygomatico-orbital complex and zygomatic arch fractures in Sao Paulo State, *Brazil, J. Oral Maxillofac. Surg.* 64 (2006) 63–67.
[2] P.S. Craighead, K. Sait, G.C. Stuart, K. Arthur, J. Nation, M. Duggan, D. Guo, Management of aggressive histologic variants of endometrial carcinoma at the Tom Baker Cancer Center between 1984 and 1994, *Gynecol. Oncol.* 77 (2000) 248–253.
[3] R. Achuthan, K. Grover, F. MacFie, Critical evaluation of the electronic surgical logbook, *BMC Med. Edu.* 6 (2006) 15.
[4] R. Glazebrook, B. Chater, P. Graham, et al., Evaluation of the ACRRM national radiology program for Australian rural and remote medical practitioners, *Rural Remote Health* 5 (2005) 349.
[5] H. Sato, S. Sato, Y.M. Wang, I. Horikoshi, Add-in macros for rapid and versatile calculation of non-compartmental pharmacokinetic parameters on Microsoft Excel spreadsheets, *Comput. Meth. Prog. Biomed.* 50 (1996) 43–52.
[6] C. Dansirikul, M. Choi, S.B. Duffull, Estimation of pharmacokinetic parameters from non-compartmental variables using Microsoft Excel, *Comput. Biol. Med.* 35 (2005) 389–403.
[7] H.A. Khan, ArrayVigil: a methodology for statistical comparison of gene signatures using segregated-one-tailed (SOT) Wilcoxon signed-rank test, *J. Mol. Biol.* 345 (2005) 645–649.
[8] H.A. Khan, ArraySolver: an algorithm for color-coded graphical display and Wilcoxon signed-rank statistics for comparing microarray gene expression data, *Comput. Func. Genom.* 5 (2004) 39–47.
[9] A.M. Brown, A spreadsheet template compatible with Microsoft Excel and iWork Numbers that returns the simultaneous confidence intervals for all pairwise differences between multiple sample means. *Comput. Meth. Prog. Biomed.* 98 (2010)76–82.
[10] H.A. Khan, SCEW: A Microsoft Excel add-in for easy creation of survival curves, *Comput. Meth. Prog. Biomed.* 83 (2006) 12–17.
[11] A. Agresti, *An Introduction to Categorical DataAnalysis (2nd ed.)*, Wiley-Interscience, (2007).
[12] J.H. Zar, *Biostatistical Analysis (5th ed.)*, Pearson Prentice-Hall, (2010).
[13] W.G. Cochran, The $\chi 2$ Test of Goodness of Fit, *Ann. Math. Statist.* 23 (1952) 315-491.

[14] P. Armitage, Tests for Linear Trends in Proportions and Frequencies, *Biometrics* 11 (1955) 375–386.
[15] D. Zelterman, *Models for Discrete Data*, Oxford University Press (2006).
[16] J.L. Fleiss, B. Levin, M.C. Paik, *Statistical Methods for Rates and Proportions (3rd ed.)*, Wiley (2003).
[17] A.E. Maxwell, *Analysing Qualitative Data*, Methuen (1961) p34.
[18] D.I. Tudehope, J. Iredell, D. Rodgers and A. Gunn, Neonatal death: grieving families, *Med. J. Aust.* 144 (1986) 290-292
[19] L. Rosenberg, D. Slone, S. Shapiro, D.W. Kaufman, P.D. Stolley, and O.S. Miettinen, Coffee drinking and myocardial infarction in young women, *Am. J. Epidemiol.*, 111 (1980) 675–681.