

AN EPIDEMIOLOGIC CALCULATOR SUPPORTING SPANISH SPEAKING VETERINARIANS IN A RESOURCE-LIMITED COMPUTING ENVIRONMENT

K Suzuki¹, MA Risso¹, G Ruiz², N Urán³, F Gutiérrez⁴, LE Mutinelli¹, ER Pons¹

¹Facultad de Ciencias Veterinarias. Universidad Nacional de La Plata.

²Facultad de Ciencias Veterinarias. Universidad Autónoma Gabriel René Moreno. Bolivia.

³Facultad de Ciencias Veterinarias. Universidad Nacional de Asunción. Paraguay.

⁴Facultad de Veterinaria. Universidad de la República. Uruguay.

ABSTRACT: *Statistical analysis is fundamental to the field of epidemiologic research. However, commercial statistical software packages are very expensive for personal use. Some free packages are available from the Internet, but researchers do not always have access to a computer for installing such software as they like. The authors developed an epidemiologic calculator, enabling sample size calculations as well as simple analyses, for Spanish speaking veterinarians in a resource-limited computing environment. There are at least four advantages to using the calculator, such as (1) no installation required, (2) no on-line connection required during the calculation process, (3) the open-source attitude, and (4) the public-domain attitude.*

KEY WORDS: epidemiology, sample size calculation, statistical analysis

CALCULADOR EPIDEMIOLÓGICO QUE ACEPTA VETERINARIOS HISPANOPARLANTES EN UN AMBIENTE INFORMÁTICO DE RECURSOS LIMITADOS

RESUMEN: *El análisis estadístico es fundamental en el campo de la investigación epidemiológica. De todas maneras, el software comercial para análisis estadístico es demasiado caro para un uso personal. Algunos programas gratuitos se consiguen en Internet, pero los investigadores no siempre tienen acceso a una computadora para instalar dicho software como quisieran. Los autores desarrollaron un Calculador Epidemiológico, permitiendo cálculos de tamaño muestral, así como análisis simples, para veterinarios hispano parlantes en un ambiente informático de recursos limitados. Tenemos entonces, al menos, cuatro ventajas por el uso del Calculador: (1) No se necesita instalación, (2) No se requiere conexión a la red durante el proceso de cálculo, (3) Una actitud de código abierto, además de, (4) Una actitud que proponga el dominio público.*

PALABRAS CLAVE: epidemiología, cálculo de tamaño muestral, análisis estadístico

Fecha de recepción: 12/09/06

Fecha de aprobación: 28/10/07

Dirección para correspondencia: Kuniaki Suzuki, Proyecto de Desarrollo Profesional Continuo para los Veterinarios del Sur (PROVETSUR). Facultad de Ciencias Veterinarias. Universidad Nacional de La Plata. C.C. 296, (B1900AVW) La Plata, ARGENTINA. Fax: +54 221 425 3276. **E-mail:** pvs@provetsur.net

INTRODUCTION

Veterinary researchers are required that not only conduct experiments, but design the study and analyse the results appropriately. Specifically, statistical analysis is fundamental to the field of epidemiologic research (1). The problem is that commercial statistical software packages are at considerable cost for personal use. Some packages are available from the Internet at no cost to Spanish speakers for research (2, 3, 4). However, researchers do not always have access to a computer to install such software as they like. Although there are some on-line statistical calculators and applets in English, not requiring the installation, reliable access to the Internet must be provided during the calculation process (5, 6). The objective of the work was to find a solution to such problems, by developing an epidemiologic calculator, enabling sample size calculations as well as simple analyses.

MATERIALS AND METHODS

One of the most commonly used spreadsheet software, Microsoft Excel (Microsoft Corporation, Redmond, USA) was selected as a platform and used for producing the calculator. Five spreadsheets for sample size calculations such as (i) sampling for a survey, (ii) sampling for detecting a disease, (iii) sampling for detecting a difference between groups, (iv) sampling for case-control studies, and (v) sampling for testing a relative risk were produced and put into the "Tamaño muestral" file prepared. Another five spreadsheets for simple analyses using 2-by-2 tables such as (vi) measuring agreement using Cohen's kappa, (vii) chi-square and McNemar tests, (viii) measuring

associations, (ix) Mantel-Haenszel method, and (x) old style Mantel-Haenszel method were also produced and put into the "Tabla de dos por dos" file prepared. Formulas used in the calculator were generated with the information from well-established publications (7, 8, 9). The statistical software Stata (StataCorp, College Station, USA) was used for proving calculation accuracy in the files produced.

RESULTS

Two Excel files with the implementations of the sample size formulas, and measures and analyses for 2-by-2 tables were created. Figure 1 and 2 present screenshots of a spreadsheet in each file. The files were provided for 10 participants in the regional postgraduate training sessions held at the Faculty of Veterinary Sciences, National University of La Plata in the year 2006 for testing operability. In conjunction with publication of the present article, the files can be downloaded at <http://provetsur.net/>.

DISCUSSION

A calculator for sampling and simple analyses was made for Spanish speaking veterinarians in a resource-limited computing environment. There are at least four advantages to using the calculator. (1) No installation required: Users have only to access to a shared-computer at school, Internet café and other public places, even if they have no right to software installation to the computer. (2) No on-line connection required during the calculation process: Users do not have to annoy at the Internet connection point or time

Figure 1. Screenshot of a sheet for estimation of difference between means in the "Tamaño muestral" file

Figura 1. Captura de la página estimar diferencias entre medias en el archivo "Tamaño muestral"

Tamaño de la muestra para estimar diferencias entre medias			
Prueba de una cola		Prueba de dos colas	
Media de grupo 1	2000	Media en grupo 1	2000
Media de grupo 2	1750	Media en grupo 2	1750
Diferencia de medias	250	Diferencia	250
Desviación estándar común (SD)	500	Desviación estándar común (SD)	500
Nivel de confianza (1-alfa)	0,95	Nivel de confianza (1-alfa/2)	0,95
Z-alfa	1,64	Z-alfa/2	1,96
Potencia (1-beta)	0,80	Potencia (1-beta)	0,8
Z-beta	0,84	Z-beta	0,84
Tamaño muestral requerido (por grupo) (n)	49,46	Tamaño muestral requerido (por grupo) (n)	62,79
Tamaño muestral redondeado (por grupo) (n)	50	Tamaño muestral redondeado (por grupo) (n)	63
Diferencia Mínima detectable			
Prueba de una cola		Prueba de dos colas	
Tamaño de la muestra (n)	300	Tamaño de la muestra (n)	300
Desviación estándar común (SD)	50	Desviación estándar común (SD)	50
Nivel de confianza (1-alfa)	0,95	Nivel de confianza (1-alfa/2)	0,95
Z-alfa	1,64	Z-alfa/2	1,96
Potencia (1-beta)	0,8	Potencia (1-beta)	0,8
Z-beta	0,84	Z-beta	0,84
Diferencia mínima detectable (d)	10,15	Diferencia mínima detectable (d)	11,44

Figure 2. Screenshot of a sheet for measurement of association in the “Tabla de dos por dos” file
 Figura 2. Captura de la página para medidas de asociación en el archivo “Tabla de dos por dos”

		Enfermedad		Nota:
		+	-	Mientras todas las medidas puedan ser calculadas, no todas tienen sentido para un estudio dado.
Exposición	+	60	20	
	-	30	55	
	Suma	90	75	
	Suma			165
Chi-cuadrado corregido por Yates				24,629
Valor-P				0,000
Medidas de asociación:				
		95	% límite de confianza	Z-alfa 1,96
		Muy bajo	Muy alto	
Prevalencia	P	0,545	0,469	0,621
Riesgo relativo	RR	2,125	1,552	2,910
Odds ratio	OR	5,500	2,804	10,789
Riesgo relativo de la población	RRpop	1,545		
Odds ratio población	ORpop	2,200		
Efecto de medidas:				
Riesgo atribuible	AR	0,397		
Fracción atribuible	AF	0,529		
Fracción atribuible estimada	AFest	0,818		
Efecto de medidas:				
Población de riesgo atribuible	PAR	0,193		
Población de la fracción atribuible	PAF	0,353		
Población estimada a la fracción atr	PAFest	0,545		

and can concentrate on calculating. (3) The open-source attitude: Users can check the process of calculations to see the formulas in each sheet. Regarding such process, the calculator can avoid becoming a sealed black box like the attitude of some commercial statistical software packages. (4) The public-domain attitude: The use of the calculator is free of charge and also available using the Calc in the OpenOffice.org software (Spanish version at <http://es.openoffice.org/>) at no cost, if users do not have access to Microsoft Excel environment. Hereafter, users manual as well as some advanced analysis methods should be provided as may be necessary.

ACKNOWLEDGEMENTS

This work was conducted as part of the project for the capacity development for improvement of livestock hygiene in the southern part of South America through regional cooperation [in Spanish: Proyecto de desarrollo profesional continuo para los veterinarios del Sur (PROVETSUR)], funded by the Japan International Cooperation Agency. Valuable and insightful comments by the reviewers are gratefully acknowledged.

BIBLIOGRAPHY

1. Dohoo I, Martin W, Stryhn H. Veterinary epidemiologic research. AVC Inc., Charlottetown (Canada), 2003, p. 706.
2. Epi Info en Español versión 3.2.2. <http://www.cica.es/epiinfo/>, accessed on 20 August 2006.

3. Thrusfield M, Ortega C, de Blas I, Noordhuizen JP, Frankena K. WIN EPISCOPE 2.0: improved epidemiological software for veterinary medicine. *Vet. Rec.* 2001, 148 (18): 567-72.
4. OPS-Xunta de Galicia. Epidat versión 3.1. <http://dxsp.sergas.es/default.asp>, accessed on 20 August 2006.
5. GraphPad Software. QuickCalcs. <http://www.graphpad.com/quickcalcs/index.cfm>, accessed on 20 August 2006.
6. AusVet. Epi Tools. <http://www.ausvet.com.au/epitools/>, accessed on 20 August 2006.
7. Dawson B, Trapp RG. Basic & clinical biostatistics. 3rd ed. McGraw-Hill, New York (USA), 2001, p. 399.
8. Rothman KJ, Greenland S. Modern epidemiology. 2nd ed. Lippincott, Philadelphia (USA), 1998, p. 738.
9. Fletcher RW, Fletcher SW. Clinical epidemiology. 4th ed. Lippincott, Philadelphia (USA), 2005, p. 252.