

# LETTER TO THE EDITOR

## Pitfalls in statistical methods

To the Editor:

We read with great interest the study of Koh et al.<sup>1</sup> and the related editorial.<sup>2</sup> Gibbons and Hodge<sup>2</sup> outlined the importance to avoid surrogate end-points as outcomes in prognostic study and stressed the importance of the number of events (and not merely of the sample size) as critical for the power of prognostic studies. However, as regards the criticism to the 95% confidence intervals (CI) for event rates, some comments might be useful. The CI from SAS package presented by Gibbons and Hodge<sup>2</sup> appear wider than those reported by Koh et al.<sup>1</sup> Gibbons and Hodge<sup>2</sup> state that the more commonly employed SAS package provides estimates of the 95% CI which are consistent with the CI based on proportions that are shown in standard statistical texts. It must be also outlined that many methods have been proposed to calculate CI for proportions and the analyzed statistical packages do not treat the subject uniformly.<sup>3</sup> The method used by Gibbons and Hodge<sup>2</sup> seems to be the Wald method, but it should be noted that this method is not recommended when small proportions are observed. Perhaps, in this case the method of Wilson is more appropriate.<sup>4</sup> The immediate *cii* command of Stata software (StataCorp. 2011. Stata Statistical Software: Release 12. College Sta-

tion, TX: StataCorp LP) is implemented with option to obtain different CIs for single proportions. The results we obtained using the Wilson option are reported in Table 1. We were able to duplicate the 95% CI reported by Gibbons and Hodge<sup>2</sup> using the Wald option, but none of the options in Stata (*binomial*, *poisson*, *exact*, *wald*, *wilson*, *agresti*, *jeffreys*) was able to duplicate the results of Koh et al.<sup>1</sup> Our analysis confirms the broad overlap between the 95% CI for different cut-off values of summed stress score in Koh's series, but also outlines that using different statistical approaches as well as different softwares may lead to different results.

Mario Petretta, MD<sup>a</sup>

Alberto Cuocolo, MD<sup>b,c</sup>

<sup>a</sup>Department of Internal Medicine,  
Cardiovascular and Immunological Sciences,  
University Federico II, Naples, Italy

<sup>b</sup>Department of Biomorphological and  
Functional Sciences,

University Federico II, Naples, Italy

<sup>c</sup>SDN Foundation, Institute of Diagnostic and  
Nuclear Development, Naples, Italy

cuocolo@unina.it

**Table 1.** Event rates by summed stress score with 95% CI

	Summed stress score			
	0	1-5	6-10	≥11
N = 176	111	16	19	30
Events = 11	3	2	2	4
Event rate (%)	2.7	12.5	10.5	13.5
CI reported by Koh (%)	1.6-5.3	8.6-19.5	8.3-11.2	2.9-18.9
CI reported by Gibbons and Hodge (%)	0-5.7	0-28.7	0-24.3	1.2-25.6
Wilson CI from Stata (%)	0.92-7.65	3.5-36.0	2.94-31.4	5.31-29.7

## References

1. Koh AS, Flores JL, Keng FY, Tan RS, Chua TS. Correlation between clinical outcomes and appropriateness grading for referral to myocardial perfusion imaging for preoperative evaluation prior to non-cardiac surgery. *J Nucl Cardiol* 2012;19:277-84.
2. Gibbons RJ, Hodge D. The limitations of small outcome studies. *J Nucl Cardiol* 2012;19:230-2.
3. Vollset SE. Confidence intervals for a binomial proportion. *Stat Med* 1993;12:809-24.
4. Newcombe RJ, Altman DG. Proportions and their differences. In: Altman DG, Machin D, Bryant TN, Gardner MJ, editors. *Statistics with confidence: Confidence intervals and statistical guidelines*. London: BMJ Book-BMA House; 2000. p. 45-56.  
doi:10.1007/s12350-012-9580-3