Commentary S1

Error reporting the test statistics and significance levels, and arguable model building

Dear Sir,

With academic and professional interests, we read the article by Ekelund et al. (1). Conducted on 443 consecutive patients with psoriasis in Sweden, this paper examined the relationship between measures of disease severity (e.g., as measured by Dermatology Life Quality Index) and associated costs in patients with plaque psoriasis.

The authors used linear regression model to examine the relationship between the Dermatology Life Quality Index (DLQI; used as continuous variable; variable name: dlqi_tot), psoriatic arthritis (dummy variable name: dum_psor) and being on systemic treatment (dummy variable name: systemic) in relation to total cost (Table III). All of these three explanatory variables were shown to be statistically significant, as described by the authors, and depicted by the *p*-values in the reduced model (Table III).

However, the *t*-values associated with the explanatory variables are too small to reach statistical significance. It is not clear from the model what the effective sample size was (which might be attributed to missing values etc.), hence the degrees of freedom (DF) could not be definitely identified. However, as the total initial sample size was 443, we may safely assume that the maximum DF would be 439 [$DF = n-(p+1)$; where n = effective sample size, and p = number of predictors (3 in this case) in the model] (2).

The *t*-values shown also do not appear to be correct as *t*-values = (b–0)/SE [b = coefficient; SE = Standard Error] (3). Based on the given coefficients (parameter estimates) and the SEs, we re-calculated the *t*-values. Now the *p*-values of these *t*-values will depend on the DF. Assuming that the DF is 439 (the maximum possible), below are the summaries of the re-calculated estimates in STATA (4), which the authors revised and agreed to [see the erratum]. Note, if the DF is less than 439 (which is possible, but never more than 439 in this case), the *t*-values would need to be even larger to achieve statistical significance.

One other important consideration in this study is the use of linear regression. The authors mentioned that they used non-parametric (Kruskal-Wallis) test for initial univariable analyses because the outcome variable (total cost) did not meet the assumption of a normal distribution (page 685). Because it was not mentioned whether the outcome variable had a conditional normal distribution, or whether the errors were normally distributed (any of these would be more accurate than outcome being non-normally distributed), justification of using linear regression is questionable (5). Non-parametric regression or bootstrap techniques are recommended in these cases which conventionally require larger sample size (6). Otherwise, in cases of normality violation, the estimates of the SE would be affected and so would be confidence intervals and significance (5).

REFERENCES

1. Ekelund M, Mallbris L, Qvitzau S, Stenberg B. A higher score on the dermatology life quality index, being on systemic treatment and having a diagnosis of psoriatic arthritis is associated with increased costs in patients with plaque psoriasis. Acta Derm Venereol 2013; 93: 684–688.
2. Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE. Linear regression. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. 2nd ed. New York Dordrecht Heidelberg London: Springer; 2012. p. 69–138.
3. Rosner, B. Fundamentals of Biostatistics. Boston, MA: Cengage Learning, Inc.; 2011.
4. StataCorp. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP; 2011.
5. Li X, Wong W, Lamoureux EL, Wong TY. Are linear regression techniques appropriate for analysis when the dependent (outcome) variable is not normally distributed? Invest Ophthalmol Vis Sci 2012; 53: 3082–3083.
6. Hawkins DM, Olive DJ. Inconsistency of resampling algorithms for high-breakdown regression estimators and a new algorithm. J Am Stat Assoc 2002; 97: 136–159.

Nazrul Islam,
Faculty of Medicine,
University of British Columbia,
Vancouver BC, V6T 1Z3, Canada.
E-mail: nazrul.islam@ubc.ca