



Nonparametric confidence intervals for sensitivity and specificity from multiple raters

Hiroyuki Saeki*

FUJIFILM RI Pharma Co., Ltd., Tokyo, Japan - sahiroyuki@fri.co.jp

Toshiro Tango

Center for Medical Statistics, Tokyo, Japan - tango@medstat.jp

Jinfang Wang

Graduate School of Science, Chiba University, Chiba, Japan - wang@math.s.chiba-u.ac.jp

In clinical trials designed to demonstrate the efficacy of a diagnostic imaging, multiple independent raters are often required to evaluate the images from the diagnostic imaging to confirm the inter-rater reliability. Therefore, sensitivities and specificities are estimated by multiple raters. However, there are few adequate methods to integrate correlated proportions among the multiple raters. In this paper, we introduce a method based on the bivariate random effects model to integrate sensitivities and specificities. In many cases we found that the nonparametric bootstrap confidence intervals outperform the corresponding Wald-type intervals. The proposed nonparametric bootstrap confidence intervals are shown to be trustworthy if there are sufficient number of observations available.

Keywords: Correlated proportions; Bivariate random effects model; Bootstrap.