Reliability of four different computerized cephalometric analysis programs: a methodological error

Sir,
We were interested to read the paper by Erkan and Gurel published in the June 2012 issue of Eur J Orthod. The authors aimed to compare the traditional method of manual cephalometric tracing with four different computerized tracing programs (Dolphin Imaging, Vistadent, Nemoceph and Quick Ceph). They used multivariate analysis of variance and Box’s and Levene’s tests, showing no statistically significant difference between manual tracing and the computerized tracing programs (Erkan et al., 2012).

The authors pointed out in their conclusions that the measurements obtained with the cephalometric analysis programs used in the study were reliable. However, they did not use any of the commonly used statistical tests (Intraclass correlation coefficient ICC or weighted kappa) to assess the reliability (Jeckel et al., 2007; Szkel and Nieto, 2007; Rothman et al., 2008). Therefore, we would like to point out that this conclusion may be misleading.

Siamak Sabour* and Elahe Vahid Dastjerdi**
*Department of Clinical Epidemiology and **Faculty of Dentistry, Shahid Beheshti University of Medical Sciences, Iran

References
Szkel M, Nieto F J 2007 Epidemiology beyond the basics, 2nd edn. Jones and Bartlett Publisher, Manhattan, NY

Reply
Sir,
We thank Dr Sabour and Dr Dastjerdi for their interest in our work. Analysis of errors, intra-observer reliability, and intra-class statistics of orthodontic cephalometric measurements were performed using Houston analyses (Houston, 1983).

In our study, intra-observer reliability analysis was performed as described by Houston (1983) and Houston et al. (1986) but was not mentioned in the manuscript. To evaluate intra-observer reliability, 10 radiographs were randomly selected. The same radiographs were then traced twice manually and digitally with each cephalometric tracing program, with a 10 day interval between evaluations. A linear correlation test was performed, and all measurements presented coefficients greater than 0.9. A measurement with a reliability coefficient greater than 0.7 is generally regarded as acceptable (Erkan et al., 2012) according to Houston (1983). Therefore, we believe that statistical analyses used in the study were appropriate. Thank you for giving us the opportunity to respond.

Mustafa Erkan
Department of Dental Clinic, Section of Orthodontics, Gulhane Military Medical Academy, Haydarpasa Research and Training Hospital, Istanbul, Turkey