LETTER TO THE EDITOR

Accuracy of linear intraoral measurements using cone beam CT and multidetector CT: methodological mistake

Dentomaxillofacial Radiology (2013) 42, 20130048. doi: 10.1259/dmfr.20130048

Cite this article as: Sabour S, Kermani H. Accuracy of linear intraoral measurements using cone beam CT and multidetector CT: methodological mistake. *Dentomaxillofac Radiol* 2013; 42: 20130048.

We were interested to read the paper by Patcas et al published in the December 2012 issue of DMFR. The authors compared the accuracy of linear bone measurements of cone beam CT (CBCT) with multidetector CT (MDCT) and validated intraoral soft-tissue measurements in CBCT.¹ They reported that radiological measurements are accurate with a mean difference from anatomical measurements of 0.14 mm (CBCT) and 0.23 mm (MDCT). However, limits of agreement for soft-tissue measurements in CBCT were smaller (−0.77 mm; 1.07 mm), indicating a slightly higher accuracy. As the authors point out in their conclusion, CBCT is slightly more reliable for linear measurements than MDCT and less affected by metal artefacts.¹ The common practice for assessing validity (accuracy) is to apply well-known statistical tests, and it is unclear why the authors did not consider employing such practice.^{2–6} Moreover, the reliability (precision) of a test is being assessed using different methods from those for validity (accuracy).^{2,3}

Reliability and validity are two completely different methodological issues in research. Sensitivity, specificity, positive predictive value, negative predictive value, likelihood ratio positive (true positive/false positive) and likelihood ratio negative (false negative/true negative) as well as the odds ratio (true results/false results—preferably greater than 50) and diagnostic accuracy are some of the tests to evaluate the validity (accuracy) of a single test compared with a gold standard.^{2–6} Reliability (repeatability or reproducibility) is assessed by different statistical tests such as Pearson's *r*, least squares and paired *t*-test, all of which can result in some of the common mistakes in reliability analysis.^{2–6} Briefly, for a quantitative variable,

intraclass correlation coefficient agreement and for qualitative variables weighted kappa should be used with caution because kappa has its own limitations too. Regarding reliability or agreement, it is good to know that, for computing the kappa value, just concordant cells are being considered, whereas discordant cells should also be taken into account to reach the correct estimation of agreement (weighted kappa).^{2–6} Two important weaknesses of the kappa value to assess agreement of a qualitative variable are as follows: it depends on the prevalence in each category and it also depends on the number of categories. So the fewer the categories, the higher the kappa value and vice versa, which can easily lead to misinterpretation.^{2–6}

It is crucial to know that statistics cannot provide a simple substitute for clinical judgment. As a take-home message, for reliability and validity analysis, appropriate tests should be applied with careful interpretation.

S Sabour

Department of Clinical Epidemiology

School of Dentistry

Shahid Beheshti University of Medical Sciences

Tehran

Islamic Republic of Iran

E-mail: s.sabour@sbmu.ac.ir

H Kermani

Department of Dentomaxillofacial Surgery

School of Dentistry

Shahid Beheshti University of Medical Sciences

Tehran

Islamic Republic of Iran

References

1. Patcas R, Markic G, Müller L, Ullrich O, Peltomäki T, Kellenberger CJ, et al. Accuracy of linear intraoral measurements using cone beam CT and multidetector CT: a tale of two CTs. *Dentomaxillofac Radiol* 2012; 41: 637–644. doi: 10.1259/dmfr/21152480.
2. Jeckel JF, Katz DL, Elmore JG, Wild DMG. *Epidemiology, biostatistics and preventive medicine*. 3rd edn. Philadelphia, PA: Elsevier; 2007.
3. Rothman JK, Greenland S, Lash TL. *Modern epidemiology*. 3rd edn. Baltimore, MD: Lippincott Williams & Wilkins; 2008.
4. Sabour S, Dastjerdi EV. Reliability of implant surgical guides based on soft-tissue models: a methodological mistake. *J Oral Implantol* 2012; 38: 805. doi: 10.1563/AAID-JOI-D-12-00176.
5. Sabour S. Reliability and repeatability of toe pressures measured with laser Doppler and portable and stationary photoplethysmography devices. *Ann Vasc Surg* 2012; 26: 1167.
6. Sabour S, Ghassemi F. The reproducibility of measurements of differential renal function in paediatric 99mTc-MAG3 renography: is this correct? *Nucl Med Commun* 2012; 33: 1311. doi: 10.1097/MNM.0b013e328359453a.