

24/7 Smart assistance: Comparison of the diagnostic performance of artificial intelligence software in the evaluation of breast ultrasound masses

Purpose: To compare the performance between BI-RADS reported (BR) by an experienced local radiologist and Koios DS software (KDS) for ultrasound (US) diagnosis of breast nodules.

Materials and Methods: We performed a retrospective cohort study of 945 female patients (1058 breast nodules), who underwent US-guided breast biopsy between February 2016 and May 2021. AIS (Koios DS software) was trained with a 400.000 images dataset and validated on a separate set of 1800 images. All the nodules were classified into ACR BI-RADS by a physician (BR) and AIS. An AIS (Koios DS software) automatically classifies breast lesions suspicious for cancer based on image data into one of four ACR BI-RADS aligned categories (benign, probably benign, suspicious or probably malignant) and also displays a continuous graphical confidence level indicator depicting where the lesion falls within its respective category and its relation to neighboring categories. The software automatically classifies the shape (round, oval, and irregular) and orientation (parallel, not parallel) of the selected lesion. Gold standard was the pathology report. PPV and NPV were contrasted with relative PPV (r PPV) and relative NPV (rNPV) for paired samples.

Results : Median age was 49 years (IQR: 41-60). Sn: BR 100%, KDS 96.1%. Sn KDS – BR: -3.8% (CI 95: -5.8%; -1.9%; p: 0.0001). Sp: BR 9.8; KDS 37.7% Sp KDS – BR: 27.9% (CI 95: 23.8%; 32%; p < 0.0001); Relative BR/KDS PPV: 0.82 (CI 95: 0.79; 0.85; p< 0.0001). PPV: BR 37%; PPV KDS 45% rNPV BR/KDS: 1.05 (CI 95: 1.03, 1.08; p = 0.0002) NPV: BR 100%; KDS 94.5%.KDS downstaged 240 nodules classified as BR-4+ by BR; 5.83% were malignant. KDS upstaged 33 BR-3 nodules suspicious (S) and probably malignant (PM), all were benign.BR had higher Sn and NPV, but KDS had higher Sp and PPV. It highlights that both KDS and radiologists are both complementary and essential to improving diagnostic accuracy.

Clinical Relevance Statement: KDS is a useful additional tool to improve the diagnostic performance of radiologists with different levels of expertise and could be used as a second opinion. Besides, it might be most useful in developing countries where the shortage of specialists represents a barrier in breast cancer screening.