Deep Learning for Identifying Benign Small Renal Masses on CT Scans
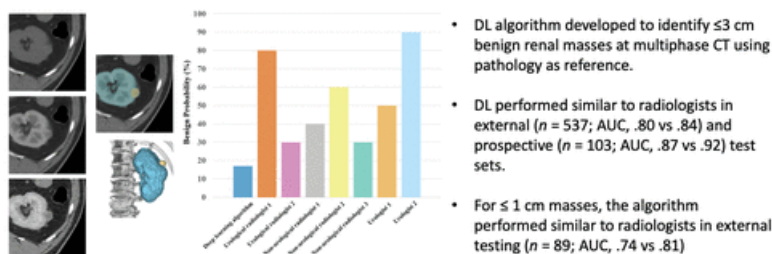

The widespread use of modern imaging has led to increased detection of small renal masses (SRMs), but earlier interventions have not reduced renal cancer-specific mortality rates, as many of these lesions are benign or indolent. Revised guidelines suggest using a size cutoff of 3 cm for T1a stage renal cell carcinoma, highlighting the importance of active surveillance or ablation as alternatives to surgery for small tumours. However, accurately distinguishing between benign and malignant SRMs remains challenging. While noninvasive imaging techniques are commonly used for characterization, approximately 14%-20% of surgically resected SRMs are pathologically benign. Invasive biopsy offers high diagnostic accuracy but is not routinely recommended. Convolutional neural networks (CNNs) have shown promise in improving diagnostic accuracy, and recent studies have demonstrated considerable accuracy in differentiating benign from malignant renal tumours based on CT images. However, limited studies have investigated the diagnostic performance for SRMs. This study published in Radiology aimed to develop and validate a deep learning algorithm for identifying small benign renal masses on contrast-enhanced multiphase CT scans and compare its diagnostic performance with that of seven clinicians.

Challenges in Distinguishing Benign from Malignant Small Renal Masses

The study evaluated 1703 patients with a single renal mass each, with a mean age of 56 years and a majority of female patients. The retrospective dataset consisted of 1063 lesions, including 874 in the training set and 189 in the internal test set. The multicenter external test set included 537 lesions, of which 12.3% were benign, with 89 lesions measuring 1 cm or less. The prospective test set comprised 103 lesions, with 13.6% being benign and 19.4% measuring 1 cm or less. Surgically resected renal masses measuring 3 cm or less in diameter at contrast-enhanced CT were included. The deep learning algorithm was developed using retrospective data from one hospital between 2009 and 2021, with patients randomly allocated to a training and internal test set ratio of 8:2. External testing was conducted between 2013 and 2021 using data from five independent hospitals, and a prospective test set was obtained between 2021 and 2022 from one hospital. Algorithm performance was assessed using the area under the receiver operating characteristic curve (AUC) and compared with the results of seven clinicians using the DeLong test.

DL Assessment of Small Renal Masses at Multiphase CT

- DL algorithm developed to identify ≤3 cm benign renal masses at multiphase CT using pathology as reference.

- DL performed similar to radiologists in external (*n* = 537; AUC, .80 vs .84) and prospective (*n* = 103; AUC, .87 vs .92) test sets.

- For ≤ 1 cm masses, the algorithm performed similar to radiologists in external testing (*n* = 89; AUC, .74 vs .81)

Dai C and Xiong Y et al. Published: May 14, 2024
https://doi.org/10.1148/radiol.232178

Radiology

Image Credit: RSNA Radiology

Performance Comparison: Deep Learning Algorithm vs. Urological Radiologists

The study focused on developing and validating a deep learning (DL) algorithm to distinguish between benign and malignant small renal masses

(SRMs) measuring 3 cm or less in diameter on contrast-enhanced CT scans. The DL algorithm's performance was compared with that of urological radiologists. The DL algorithm was trained using retrospective data from one hospital and evaluated on both internal and external test sets. Performance was assessed using the area under the receiver operating characteristic curve (AUC). In addition, the DL algorithm's performance was compared with that of seven urological radiologists.

Results showed that the DL algorithm demonstrated favourable performance, with AUCs comparable to those of urological radiologists in both the external and prospective test sets. Notably, the DL algorithm performed similarly well in identifying subcentimeter (≤1 cm) renal masses, a challenging task even for experienced clinicians. The DL algorithm performance was comparable with that of urological radiologists: for the external test set, AUC was 0.80 (95% CI: 0.75, 0.85) versus 0.84 (95% CI: 0.78, 0.88) (P = .61); for the prospective test set, AUC was 0.87 (95% CI: 0.79, 0.93) versus 0.92 (95% CI: 0.86, 0.96) (P = .70). For subcentimeter lesions in the external test set, the algorithm and urological radiologists had similar AUC of 0.74 (95% CI: 0.63, 0.83) and 0.81 (95% CI: 0.68, 0.92) (P = .78), respectively.

Role of Deep Learning in Renal Mass Characterization

The study acknowledged the limitations of using a dataset consisting solely of surgically resected cases, which may introduce bias towards malignancy. However, an adaptive sampling method was employed during model training to address this imbalance. Additionally, the DL algorithm's performance was found to be slightly lower on external datasets compared to internal ones, likely due to differences in scanning protocols.

Overall, the study concluded that the DL algorithm could effectively assist clinicians, particularly in resource-poor settings where expert radiologists may not be readily available. Furthermore, the algorithm could potentially reduce unnecessary surgical interventions for benign lesions by facilitating active surveillance for small renal masses.

Source: RSNA Radiology

Image Credit: iStock

Published on : Thu, 16 May 2024