
Nodule risk estimation in screening CT: computer model vs. human observers



A new study sought to compare human observers to a mathematically derived computer model for differentiation between malignant and benign pulmonary nodules detected on baseline screening computed tomography (CT) scans. The study findings show that the computer model (PanCan) and human observers perform equivalent for differentiating malignant from randomly selected benign nodules, confirming the high potential of computer models for nodule risk estimation in population-based screening studies.

"Human observers, however, significantly outperform the PanCan model for differentiating malignant from size-matched screen-detected benign nodules suggesting that integration of additional morphological characteristics, such as pleural retraction and perinodular lung parenchyma distortion, used by the human observers is very likely to lead to further improvement of computer-based risk prediction models," says the study published in the journal PLoS One.

Low-dose CT studies of the lung — obtained within a lung cancer screening programme — are characterised by the discrepancy between the high prevalence of pulmonary nodules and the relatively low incidence of actual lung cancers. A prospective estimation of which nodules are of high risk to represent or develop into a malignancy, and therefore requiring more close follow-up or intense diagnostic work-up as opposed to nodules that are of very low risk, is of crucial importance to make screening programmes efficient for a number of reasons including burden of radiation dose, psychological load of the screening subjects and financial expenses.

Lung-RADS and similar recommendations have in common that they use nodule type, nodule size and growth rate to select nodules that are malignant or at risk to develop into a malignancy. A number of other risk prediction models rather focus on the selection of individuals being at risk for developing a lung malignancy or estimate the malignancy probability using both, clinical factors and nodule characteristics. In the current study, researchers compared the PanCan model with human observer performance for the prediction of malignancy risk of screen-detected nodules. The study cohort consisted of 300 chest CT scans from the Danish Lung Cancer Screening Trial (DLCST). It included all scans with proven malignancies ($n = 62$) and two subsets of randomly selected baseline scans with benign nodules of all sizes ($n = 120$) and matched in size to the cancers, respectively ($n = 118$). Eleven observers and the PanCan model assigned a malignancy probability score to each nodule. Seven observers assessed morphological nodule characteristics using a predefined list.

Based on the results, performances of the model and observers were equivalent (AUC 0.932 vs. 0.910) for risk-assessment of malignant and benign nodules of all sizes. However, human readers performed superior to the computer model for differentiating malignant nodules from size-matched benign nodules (AUC 0.819 vs. 0.706). Large variations between observers were seen for ROC areas and ranges of risk scores.

The researchers say there is another way to further improve the model's performance — by including nodule growth between scans obtained at different time points. Several studies have shown that lesion growth over time is the most important and powerful predictor of nodule malignancy.

However, there is an inherent discrepancy in the process of risk estimation of a logical but rigid mathematical model and the intuitive but variable visual analysis of human observers. "Whereas a direct comparison of the scores does not lead to meaningful conclusions, the ROC statistical analysis we used sufficiently considered the relative distribution of the scores and therefore allowed for comparing the performances of observers and the PanCan model," the researchers explain.

Source: [PLoS One](#)
Image Credit: Pixabay

Published on : Wed, 22 Nov 2017