Comparing AI with Retrospectively Reconstructed Lung-RADS for Lung Cancer Detection in Low-Dose CT Screening

SUBMITTED MANUSCRIPT - ABSTRACT

Purpose: To compare an artificial intelligence (AI)-based lung cancer likelihood model with retrospectively reconstructed Lung-RADS version 1.0 for detecting lung cancer on low-dose CT (LDCT) screening examinations.

Methods

The analysis included baseline LDCT examinations from a random subset of the U.S. National Lung Screening Trial (US-NLST; 2,342 examinations, 40 one-year cancers) and the Pamplona International Early Lung Cancer Action Program (P-IELCAP; 3,709 examinations, 27 one-year cancers). AI assigned examinations to low, intermediate, high, or highest likelihood categories. Lung-RADS categories were reconstructed from radiologist-documented nodule information. PanCan was assessed as a secondary comparator. The primary outcome was lung cancer diagnosed within one year.

Results

Mean age was 61.6 years in US-NLST and 56.6 years in P-IELCAP. The high and highest AI categories contained 33 of 40 cancers (82.5%) in US-NLST and 20 of 27 (74.1%) in P-IELCAP. Positive predictive values were 13.4% (33/246) and 9.7% (20/207), compared with 8.0% (30/376) and 5.1% (20/390) for Lung-RADS category 3 or higher (P<0.001 for both comparisons). The low-likelihood AI category contained one cancer in each cohort, corresponding to observed cancer rates of 0.09% and 0.04%; rates for Lung-RADS categories 1-2 were 0.51% and 0.22%. Across categories, PanCan had higher positive predictive value but lower sensitivity than AI.

Conclusion

Compared with retrospectively reconstructed Lung-RADS, AI concentrated more one-year cancers in the high-likelihood categories and assigned fewer cancers to the low-likelihood category. Prospective studies are needed to establish the clinical utility of this approach.

Keywords: Lung cancer screening; low-dose CT; artificial intelligence; Lung-RADS

Abstract prepared from the submitted manuscript. The full manuscript is not provided for download.