Improving Radiologists' Recommendations With Computer-Aided Diagnosis for Management of Small Nodules Detected by CT[1]

Feng Li, Qiang Li, Roger Engelmann, Masahito Aoyama, Shusuke Sone, Heber MacMahon, Kunio Doi

Rationale and Objectives. To evaluate how computer-aided diagnosis (CAD) can improve radiologists' recommendations for management of possible early lung cancers on CT.

Materials and Methods. Twenty-eight lung cancers and 28 benign lesions were employed. Each group of 28 lesions was classified into subgroups of two sizes (9 between 6 and 10 mm and 19 between 11 and 20 mm) and three patterns (8 with pure ground glass opacity [GGO], 12 with mixed GGO and 8 solid lesions). Sixteen radiologists participated in the observer study, first without and then with CAD. Radiologists' recommendations, including (1) follow-up in 12 months, (2) in 6 months, (3) in 3 months, or (4) biopsy, were compared at three levels of their malignancy probability ratings (low: 1%–33%; medium: 34%–66%; high: 67%–99%) for 896 observations (56 lesions by the 16 radiologists) in the two size subgroups and three patterns.

Results. The number of recommendations changed by radiologists by use of CAD was 163 (18%) among all 896 observations. Among these changed recommendations, the fraction showing a beneficial effect from CAD was 68% (111/163), and the fraction showing a beneficial effect regarding biopsy recommendations was 69% (48/70). With CAD, the radiologists' performance regarding biopsy recommendations was significantly improved for 43 lung cancers (31 changed to biopsy versus 12 changed away from biopsy; $P = .003$) and was also improved for 27 benign lesions (10 changed to biopsy versus 17 changed away from biopsy; $P = .18$). Most of the cancers with improved recommendations were solid lesions or mixed GGO and relatively large.

Conclusion. CAD has the potential to improve the appropriateness of radiologists' recommendations for small malignant and benign lesions on CT scans.

Key Words. Lung neoplasms, CT; Computer diagnostic aid; Lung, module.

Acad Radiol 2006; 13:943–950

[1] From the Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, The University of Chicago, Chicago, IL (F.L., Q.L., R.E., H.M., K.D.); Department of Intelligent Systems, Faculty of Information Sciences, Hiroshima City University, Hiroshima, Japan (M.A.); and JA Nagano Chu-Nanshin Regional Cancer Screening Centre, Ikeda, Nagano, Japan (S.S.). Received December 13, 2005; accepted April 17, 2006. This work was partly supported by USPHS grant CA62625. H. MacMahon and K. Doi are shareholders of R2 Technology, Inc., Los Altos, CA. K. Doi is a shareholder of Deus Technology, Inc., Rockville, MD. CAD technologies developed in the Kurt Rossmann Laboratories have been licensed to companies including R2 Technology, Deus Technologies, Riverain Medical Group, Mitsubishi Apace Software Co., Median Technologies, General Electric Corporation, and Toshiba Corporation. It is the policy of The University of Chicago that investigators disclose publicly actual or potential significant financial interests that may appear to be affected by research activities.**Address correspondence to**: F.L. e-mail: feng@uchicago.edu

doi:10.1016/j.acra.2006.04.010

Among diagnostic imaging modalities, computed tomography (CT) has the highest sensitivity for detection of small pulmonary lesions. However, it is difficult for radiologists to correctly distinguish cancers from noncancerous lesions (false positives) and to make appropriate and consistent recommendations management of patients with suspicious lesions. On the one hand, a large number of false positives will lead to unnecessary patient anxiety and will increase the increased economic costs and radiation exposure. A high rate of false positives can also lead to unnecessary investigation such as CT scans, biopsy, and even surgery. On the other hand, in the case of lung cancers (true positives), if radiologists fail to make an appropriate recommendation such as biopsy or surgery, the patients may miss an opportunity for cure.

The Food and Drug Administration has approved the clinical use of some computer-aided diagnosis (CAD) detection systems in screening for clinical use, especially for breast cancer screening on mammography in the United States. Gur et al (1) reported that the introduction of detection CAD into a large clinical practice (115,571 screening mammograms) was not associated with statically significant changes in both recall and breast cancer detection rates. Commercially available detection CAD systems show marks, including true positives (cancers) and false positives (noncancerous lesions also anatomic structures), on each whole image (1–3). Recently, automatic classification CAD schemes for distinction of malignant and benign lesions have been developed in some universities (4–8) that show an estimated likelihood of malignancy for each segmented lesion based on its image features. Some observer studies using mammograms reported that classification CAD had a beneficial effect for radiologists' diagnostic accuracy for classifying malignant and benign breast masses and their recommendations regarding biopsy (5,6).

It is important that a larger database, including large number of lesions and a variety of lesion patterns, be used for developing classification CAD. The thin-section CT database for developing our classification CAD scheme used in this study comprised follow-up exams obtained from a 3-year CT lung cancer screening program (17,892 examinations). The database included 61 primary lung cancers (size range 6–19 mm; mean 12 mm) and 183 benign nodules (size range 3–20 mm; mean 7 mm) with three different patterns (8,9). We have reported (8) that our CAD scheme has the potential to improve radiologists' diagnostic accuracy for lesion classification and also to improve radiologists' recommendations in an observer study. The data analysis in the previous report (8) was independently calculated for 16 observers, and the radiologists' recommendations were improved by increasing the number of biopsy recommendations for actual early cancers (statistically significant) and by reducing the number for actual benign ones (not significant) in an observer study. The current study used the same data from the same observer test as used previously (8). Our purpose in this study was to evaluate further how CAD can assist radiologists in their recommendation management of possible early lung cancers that have different sizes and patterns.

MATERIALS AND METHODS

Institutional review board approval and informed observer consent were obtained.

Database

Our database was obtained as part of an annual 3-year CT screening for lung cancer in a general population in Nagano, Japan (8,9), which included 59 patients (27 men, 32 women, mean age 64.6 years) with 61 primary small lung cancers (mean size 12.3 mm; size range 6–20 mm; 18 nodules with pure ground glass opacity [GGO]; 28 with mixed GGO; and 15 with solid opacity), and 169 patients (99 men, 70 women, mean age 61.6 years) with 183 benign lesions (mean size 7.2 mm; size range 3–20 mm; 12 with pure GGO, 30 with mixed GGO, and 141 with solid opacity). All patients gave informed consent. All cancers were confirmed by surgery, and benign lesions were confirmed by surgery or follow-up (resolved or no change for 2 years or more). The mean size (average length and width) of each nodule was recorded by one radiologist (F.L.). The three types of patterns of these lesions, including pure GGO, mixed GGO, and solid opacity, were viewed independently and grouped by three radiologists (F.L. among them) without knowledge of the final diagnosis, and then a consensus was reached through discussion. Thin-section CT scans were performed on a helical scanner (CT HiSpeed Advantage, GE, Milwaukee, WI) with a standard tube current (200 mA) to cover the entire lesion, with 1-mm collimation and a bone reconstruction algorithm with a 0.5-mm interval.

CAD

With our CAD scheme, the nodules were segmented automatically by use a dynamic programming technique.