Lung Image Database Consortium

Subcommittee on Inclusion Criteria (2)

Report #6

Summary of Key Findings from Reports #1-5

A. Types of studies

The Lung Image Database should not necessarily be limited to screening CT studies and diagnostic CT scans obtained as follow-up to screening CT studies. In other words, the Database could contain diagnostic CT scans acquired from patients who did not undergo CT screening. Retrospective CT scans for inclusion in the Database may have been obtained from either single slice or multislice scanners, but a maximum collimation of 5 mm will be imposed. Lesion-targeted scans will be included in the database as well.

B. Prospective vs. retrospective cases

The collection of retrospective cases would significantly increase the pool of cases available for inclusion in the Database. The Database will benefit from both retrospective and prospective cases.

C. Nodule size criterion

The spatial locations (e.g., the centers) of <u>all</u> lesions (including scars) will be recorded for each scan in the Database. Only nodules greater than 3 mm will be subjected to (1) observation for the purpose of collecting follow-up scans and eventual pathologic truth and (2) the type of analyses required for further characterization of the lesion (e.g., outlines by one or several experts). A nodule originally thought to be greater than 3 mm that eventually measures slightly less than 3 mm through tools on the interface will <u>not</u> be disqualified.

D. Nodule definition

Consensus was reached on use of the term "nodule" instead of "focal opacity." The term "focal abnormality" was adopted as the more generic category of lesions that contains nodules as a subclass. To characterize lesions identified as nodules, a "nodule class" and "margin descriptors" will be associated with each nodule:

nodule class	margin descriptors
• solid	• smooth
• part-solid	• spiculated
• nonsolid	• poorly defined

The margin descriptors should be rated (e.g., on a 5-point scale) to capture a spectrum of appearances. Additional descriptors are being considered as well. Borrowing from Henschke, *et al*. (AJR 2002; 178:1053–1057), "solid" nodules may be distinguished from "subsolid" nodules in that the former class of nodule "completely obscures the entire lung parenchyma within it whereas the latter does not." Moreover, subsolid nodules may be classified as either "part-solid" nodules, in which "patches of parenchyma…are completely obscured," or "nonsolid" nodules, in which no parenchymal obscuration is present. It was agreed that the database should include any nodule that meets the minimum size requirement regardless of etiology or potential for malignancy. Cases with metastatic disease will be included.

E. Number of nodules

A limit of 6 primary lung cancers and benign lesions combined (of sufficient size) will be imposed for a scan to qualify for inclusion in the Database. The presence of additional lesions less than 3 mm will not affect the disposition of the scan. From a practical perspective, the limit of 6 was recognized as a guide: upon closer scrutiny by the panel of radiologists, some additional nodules greater than 3 mm may be identified, while some lesions originally thought to be less than 3 mm may be determined through tools on the interface to be greater than 3 mm; the scan will not be disqualified from the Database for these reasons. Furthermore, a subsequent scan may reveal more than 6 nodules greater than 3 mm despite the presence of fewer than 6 nodules in the previous scan; the scans will not be disqualified for this reason. With regard to metastatic nodules, the limit of 6 will be imposed initially; however, experience with case selection may show this limit to be too low.

F. Types of abnormalities

The notion of an "abnormal case" is not a binary concept but rather a spectrum. The Database should contain cases that capture that spectrum. Cases will be selected predominantly for their ability to populate the Database consistent with the guidelines to be established by the LIDC; however, at least some component of the Database should comprise a series of consecutive cases. "True" normals will likely be required.

G. Exclusion criteria

The one broad category of case that should be excluded from the Database is streak or motion artifact cases such that the diagnostic utility of the case is adversely affected.

H. "Truth"

A standardized protocol for establishing "truth" in the CT scans will be established. In this context, "truth" represents the spatial location of abnormalities within the scan. The spatial location of each lesion (even those generally regarded as being of no clinical significance) must be indicated, inventoried, and stored.