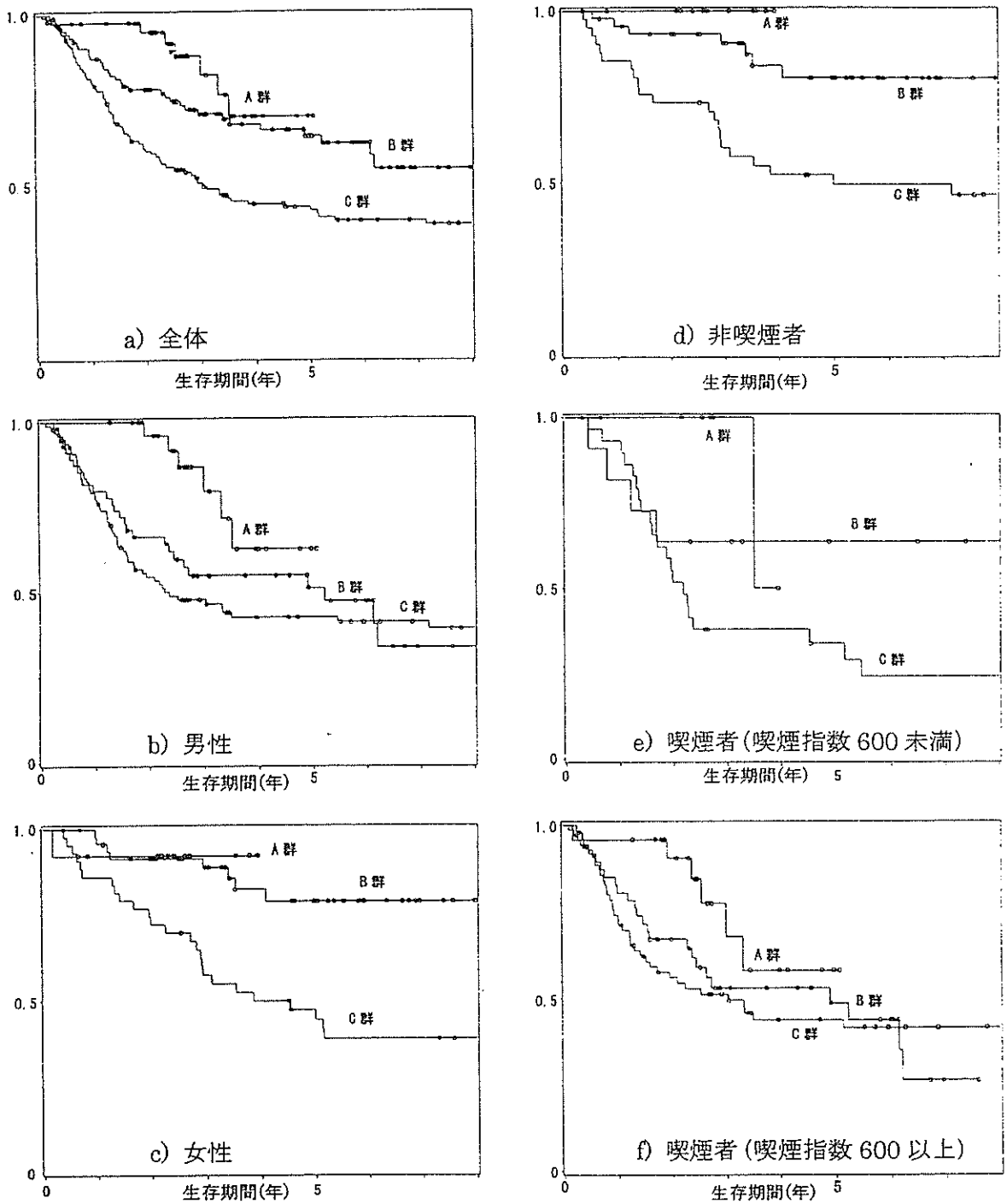



図１．検診手法別発見肺癌の男女別・喫煙指数別生存率曲線

A群：CT 検診発見群、B群：従来型検診をスクリーニング法とし、精検として CT を行った群、C群：従来型検診をスクリーニング法とし、精検として直接単純撮影や断層撮影を行った群

読影セミナー

# 肺がん検診における読影のピットホール

楠 洋子[1)]、中山富雄[1)]、鈴木隆一郎[1)]、有澤 淳[1)]、黒田知純[2)]
[1)] 大阪府立成人病センター
[2)] 大阪がん予防検診センター

（本稿は、第8回読影セミナーで発表した内容である）

自施設が CT 検診に携わってきた経験からの工夫と読影上の留意点などを述べる。新たに CT 検診を開始される方々への多少の参考になれば幸いである。

【検診成績】

まず我々の CT 検診の成績から述べる。撮影方法はシングルスライス CT による検診で、管電流 25mA、スライス厚 10mm で、住民を対象に実施している。受診者数は約4年間でのべ 9,296 人（男性 5,803 人、女性 3,493 人）で、原発性発見肺がんは 49 例（発見率 0.53%）、男性 35 人（0.60%）、女性 14 人（0.40%）であった。病期 I 期率は男女でそれぞれ 72%と 81%で、切除率は 92%と 75%であった（Table 1）。初回検診群と経年検診群（1年前に受診歴のある者）で比較してみると、精検率は初回で初回受診者 5,669 人（受診者の 61.0%）中の肺がんは 40 例（0.71%）のべ 43 部位で、IA 期率は 70.3%であった。経年受診者 3,627 人（39.0%)中の発見肺がんは 9 例（0.25%）と初回例の3分の1であった。これは経年受診者が 40%に満たないぐらいまだ少ないためと、経年受診者中には非喫煙者が大半を占めていたことにも起因すると思われた。しかし IA 期は 7 例（77.8%）と初回受診者より上回っていた(Table 2)。発見肺がんの内容をみると、初回発見例には非喫煙者が半数近く占めており、これらの組織型は全例腺がんであった。経年発見例は全例喫煙者であった。組織型では、初回発見例では腺がんが大多数（85%）を占め、そのうちの 77%が野口 type A か type B で占められていた。経年発見例は、腺がん4例（うち1例の type A を除いて type D 以上であった）、扁平上皮がん2例、大細胞がん2例、小細胞がん1例で、進行の速いがんが大多数（89%）を占めた。進行の遅いがんは初回検診時に網羅されたことも原因の一つであろう(Table 3)。

【読影の手順と工夫】

自施設では CT 検診の読影医師として携わっているのが 2 人しかおらず、当初は二重読影後に問題症例を話し合って最終判定をしていたが、時間的制約もあり独立して読影した後は、後で診たものが責任を持って最終判定を行っている。判定が異なる場合は強い方の決定に従うことが多い。

検診の常の基本である比較読影は同一 CRT 画面上のシネモードで同期させて行えるようになっている（Fig.1）。また現像した過去の CT フィルムがあれば併用している。やや大きな陰影であれば、必要時に過去の間接 XP フィルムも参考にしている。CT 読影時には過去情報や過去フィルム、磁気媒体画像など全て手元に揃えられており容易に比較ができる方針を採っている。これらの資料は全

て院内の担当保健師が検診受診者のデータベースを検索して準備を整えている。さらに判定に迷う場合は現像後にフィルムで再読を行って最終判定を決め、精検の精度を高めるようにしている。

【経年発見例の画像】

新病変の3例を除いて初回検診時に誤判定していたのは6例と多く、うち3例は初回検診時にも明らかな結節影として認められたが、共に他部位に陳旧性結核の陰影があり一元的な器質化陰影と誤判定した。この3例のうち2例が診断時すでにIIIB期、IV期であった。IV期の症例は小細胞癌であり、前々年度のCT所見では無所見であった。複数回以前のCT所見との比較を怠ったために生じた判例であった。残りの2例はその部をチェックしながら血管の一部と判定した陰影であった。最後の1例は第1肋骨のpartial volume effectと誤判定したGGO所見であった。以下に実際の症例を示しながら反省点や留意点を述べる。

【読影時の留意事項】

1. 撮影条件による陰影の性状の差に注意（Fig.2）：
   検診CT画像では、充実性陰影も淡く見えるため、疑いのある微小陰影は必ずTSCTで性状を確認してフォローの方針を決定する。
2. 血管陰影との鑑別（Fig. 3）：
   血管陰影との鑑別は困難なことが多く初回時には注意が必要である。誤判定した症例は初期の頃の症例ばかりではないため、読影の困難さを改めて感じた。発見時に進行がんでないように初回時には神経質な位の読影が必要である。精検率が高くなるという課題は残るが経年になれば精検率は下降する。判定に不安な時は数ヶ月後にTSCTで確認しておくのも一手である。少しでも悪性を否定しきれないときには定期健診よりもTSCTによる経過観察が有効である。精検受診者への丁寧な説明も大事である。
3. Partial volume effectなどとの誤認（Fig. 4）：
   4-a.は第1肋骨先端のpartial volume effectと誤認し初回時の判定はBとした例である。経年で受診されたので増大した陰影をチェックできた。重喫煙者であり、このようなGGO陰影でも短期間に増大してくる場合があり注意が必要である。
   4-b.は胸膜直下の陰影でpartial volume effectか、局所的な胸膜肥厚かとの判断で初回時の判定はCとした。次年度には陰影の増大が見られて精検を施行したが結局は腺がんtype Cでstage IIIBになっていた。*p*3のT3であり、胸膜直下の陰影には*p*因子が進むので注意を要する部位である。
4. 前年度C判定根拠の確認—比較読影を怠った例（Fig. 5）：
   検診CT所見に石灰化を思わせる陰影が認められた。数回の経年受診者であり、前年度のC判定との情報をみてこの陰影と誤解したことによる誤判定例である。確定診断がついてから確認すると、その前年度

（2000 年度）にはこの陰影はかろうじてわかるような淡い陰影があったので 2001 年度はすでに増大陰影であり、その時に比較読影を怠った。他部位には陳旧性結核の陰影があり、以前より C 判定であったことも勘違いの因子になったと思われる。次年度（2002 年度）の検診で陰影の増大が認められ、精検により小細胞癌と診断したが、すでに stage IV (骨転移)に進行していた。C 判定の根拠はできる限り確認する必要があると痛感した。

5. 陳旧性陰影（特に陳旧性結核陰影）には注意：

これは前 4 項に準じるが、さらには陳旧性陰影の中に紛れて誤認される場合がある。経年で検診を受診すれば stage IA で発見される可能性があるが、住民検診では必ず毎年受診するとは限らない。隔年後の検診では非 IA で発見された症例は少なくない。

6. 新病変は小さくてもチェックが必要：

どんなに小さくても新病変はチェックする。検診 CT から精検の対象とする大きさ Ø5mm 以下の点状の陰影でも新陰影であれば TSCT での確認後、その性状に応じたフォローが必要である。

—精査後の問題—

7. 長期フォロー例における留意事項（Fig. 6）：

微小陰影を TSCT で長期フォローしている際、時には肺全体の CT や、少なくとも直接 X 線撮影により他部位の新病変などをチェックする必要がある。症例は左肺の 5mm の陰影をフォローして開胸生検に至った時の staging CT で発見された陰影である。前年度と前前年度の検診時には 3mm の瘢痕と判定していた陰影でフォローの対象外としており、左肺の陰影のみ目標としていたが急速に増大してきた。確定診断時には stage IIIA になっていた。

8. その他：

- 精査にて確定診断がつかないとき、少しでも悪性を否定しきれなければ、定期健診に回すより、その性状により必要な期間内は医療によるフォローが有効である。
- 被爆の問題があるので、陰影によってはフォロー途中で検診条件に変更して比較できる可能性もあり、考慮する余地はある。

【種類による検診の違い】

1. 住民検診
   - 経年受診で進行がんが防げる可能性があるが、住民検診では他の検診体制とは異なり受診者は毎年受診するとは限らないジレンマがある。
   - ゆえに少しでも疑わしい陰影の場合には一度は TSCT でフォローするほうが望ましい。
   - 精検受診の勧奨に手数がかかる（保健師の負担が大きい）。
2. 職域検診
   - 経年受診率が高い。
   - 住民検診に比べ年齢層がやや低い。

・ 定年後の長期追跡が困難である。

3. ドック検診
   - 経年受診率が高い。
   - 精検率が高い（self selection bias の最たるものであり、勧奨する手間が省ける）。

【住民 CT 検診におけるジレンマ】

1. 検診受診や精検受診への勧奨は、各自治体の担当保健師 CT や CT 検診についての知識をもって説明に当たることが要求される。
2. 自治体の担当事務官や保健師は短期間で交代するので、そのつど新しい担当者へ CT 検診についての知識を共用してもらえるように教育が必要である。
2. 検診結果により、受診者は心配して電話で問い合わせをしてくることがしばしばある。自治体の保健師や当院の担当保健師および担当医である我々が説明にあたる。これら３～４者の間での往復連絡と説明に結構時間が割かれる。
3. 反対に“症状もないので精検を拒否する”人への精検勧奨も同じルートで説明を要求される場合がある。
4. 精検決定時の情報（現治療歴など）や精検準備時の情報（造影検査時の必要な基本検診での検血データの問い合わせなど）のために問診表に記載されていない情報収集を必要とする場合があり、検診機関と自治体との連絡を頻回に行わなければいけない。

などなど、住民検診においては、各施設の保健師さん達の理解と努力にかかわることが多いことが痛感される。

Table 1. CTスクリーニングの成績
(1998.10 - 2003.3*)

| | 男性 | 女性 |
|---|---|---|
| 受診者数 | 5,803 | 3,493 |
| 精検率 (%) | 8.2 | 5.9 |
| 精検受診率 (%) | 90.4 | |
| 発見肺癌数 | 35 | 14 |
| 対10万人比 | 603 | 401 |
| 病期Ⅰ期率 (%) | 72 | 81 |
| 切除率 (%) | 92 | 75 |

Table 2. 初回と経年受診による発見率

| | 初回受診 | 経年受診* |
|---|---|---|
| 受診者数 | 5,669 | 3,627 |
| 原発性肺がん | 40 (0.71) | 9 (0.25) |
| 対10万人比 | 706 | 248 |
| 病期ⅠA期率 (%) | 70.3 | 77.8 |

* 50% が非喫煙者　　( ): %

Table 3. 初回と経年の発見肺がん

| | 初回発見 n=40 | 経年発見 n=9 |
|---|---|---|
| 非喫煙者 | 17 (all AD, 男性 4) | 0 |
| 喫煙者 | 23 | 9 (女性 1) |
| 喫煙指数 | 960 (100-2040) | 1016 (340-3240) |
| < 600 | 4 | 3 |
| 600 ≤ | 19 | 6 |
| 組織型 | | |
| AD | 34 (85%) * | 4 (44%) ** |
| SQ | 3 | 2 |
| LA | 1 | 2 |
| SM | 1 | 1 |
| other | 1 | |

* Smoker/non-smoker: 17/17, Type A & Type B: 77%

** Type A: 1 pts

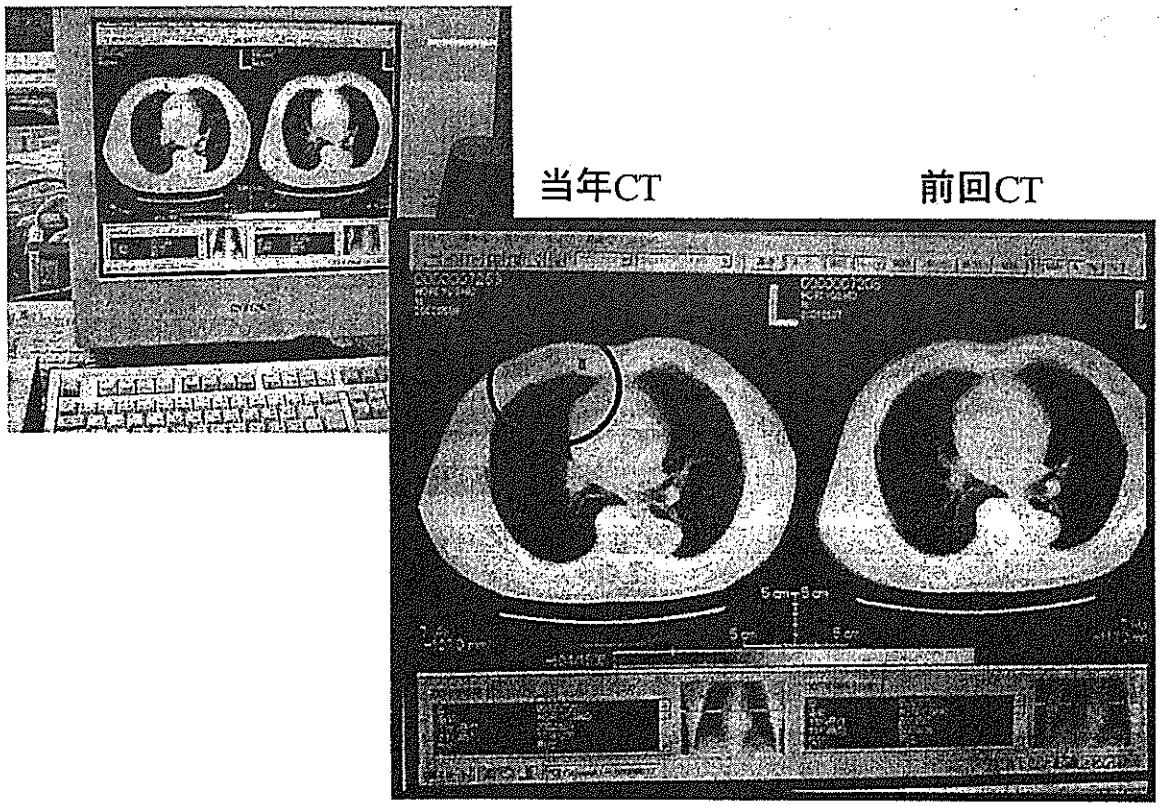
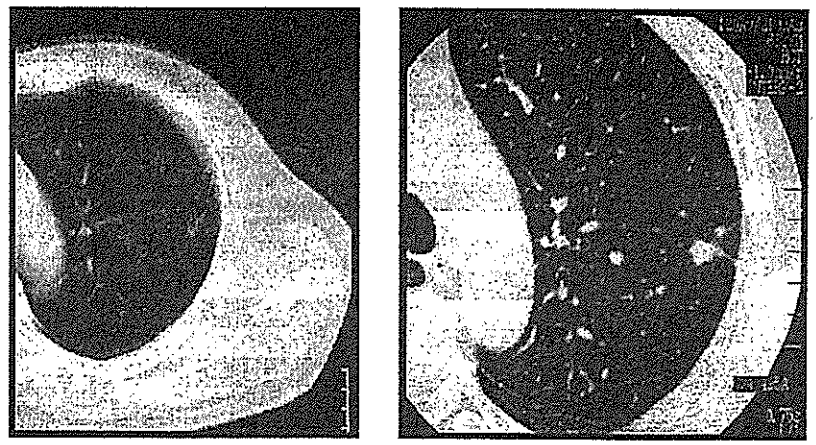



Fig.1 CRT比較読影



検診CT 判定 E1
淡く見える

TSCTで充実性

Fig.2 検診CTでの淡い陰影
AD type D, stage IA

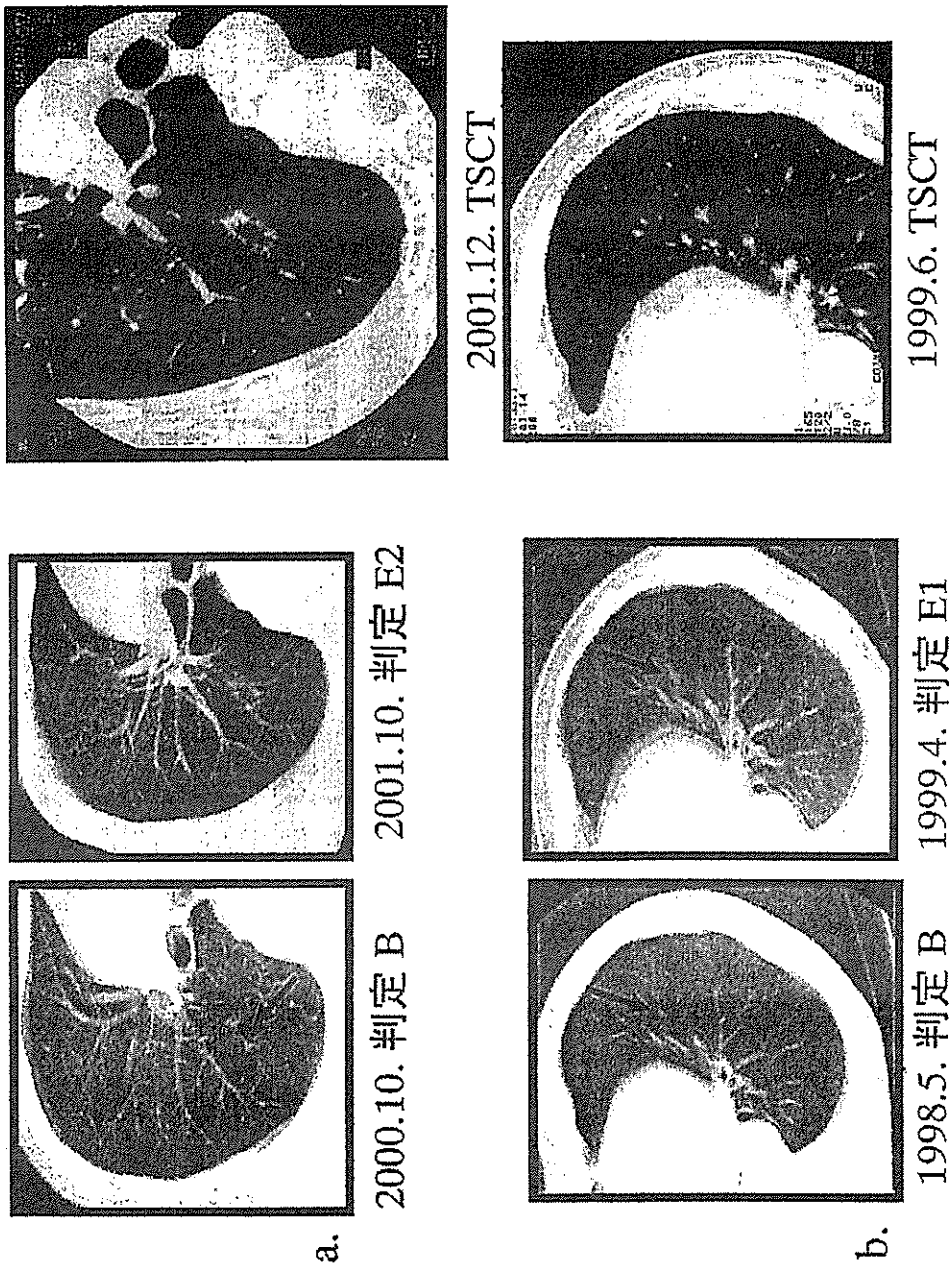



Fig.3 血管と紛らわしい陰影

a: AD type C, stage IA
b: SQ, stage IA

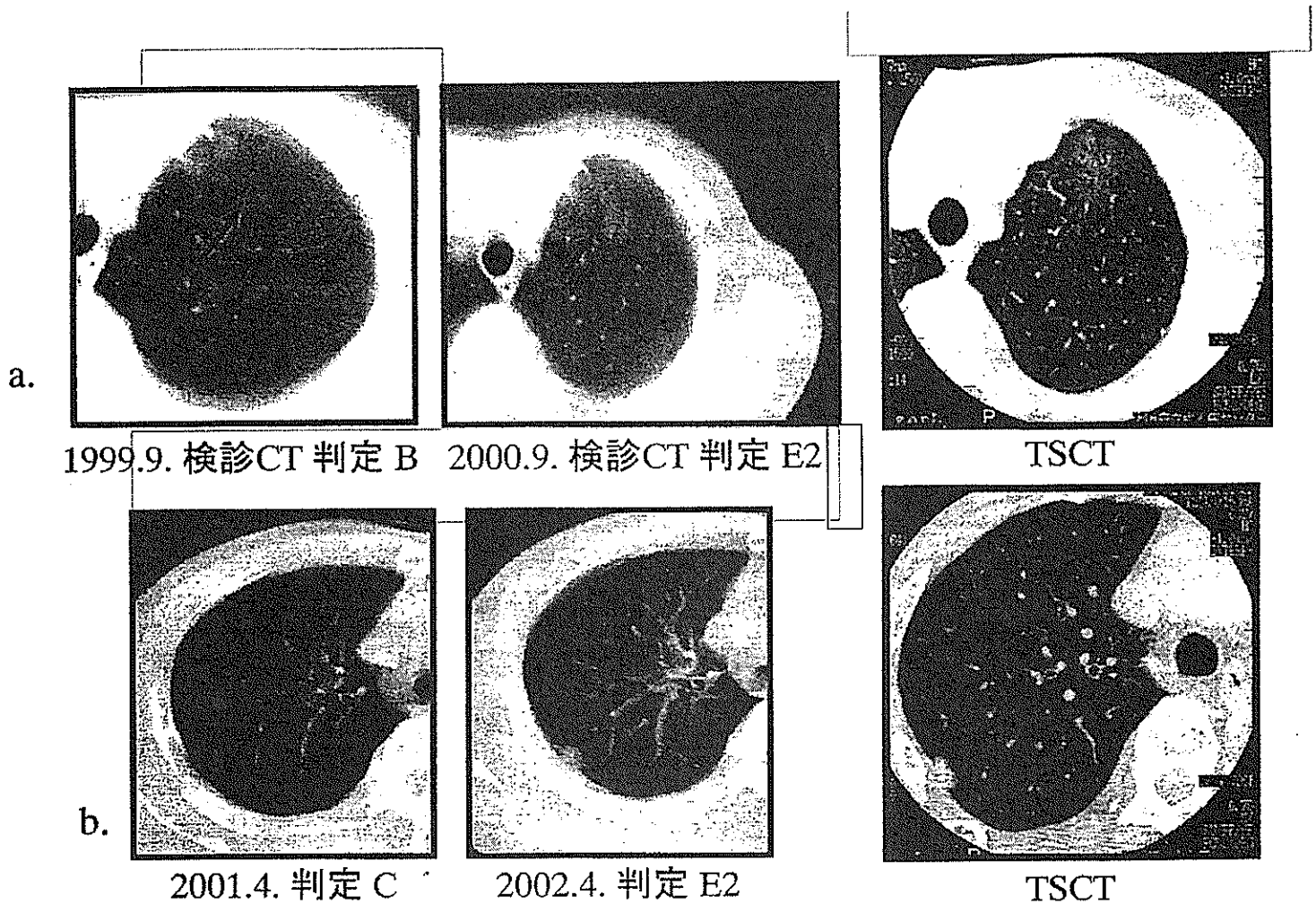
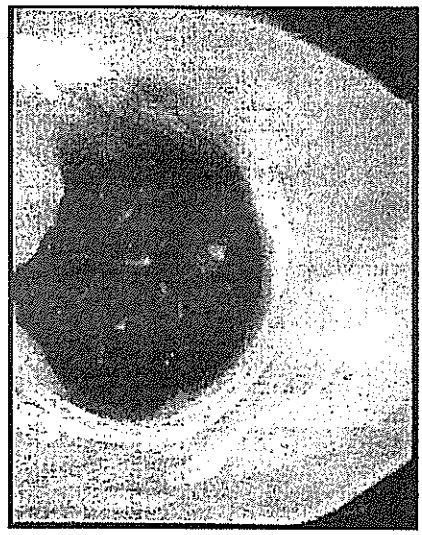


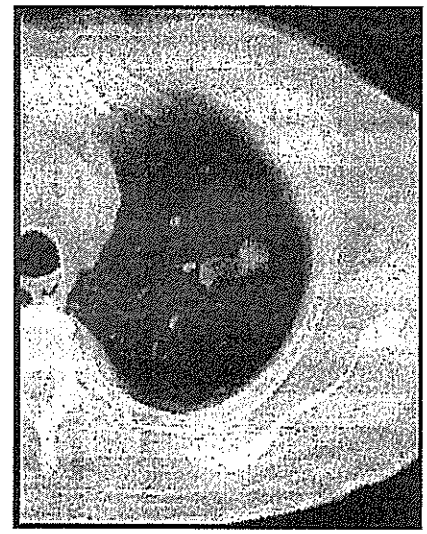

Fig. 4 Partial volume effect などとの誤認

a: 肋骨先端のpartial volume effect と誤認（AD type C, stage IA), doubling time: 145日

b: 胸膜肥厚と誤認（AD type C, stage IIIB）



2001.6. 判定 C



2002.6. 判定 E2

前年度もＣ判定
他部位に陳旧性結核陰影あり
矢印部位を石灰化と誤認

Fig. 5 比較読影を怠った症例
SM, stage IV (bone)

1999.4. 検診CT 判定C

1999.11. 通常 CT

Fig. 6 対側 or 他部位の見逃し
（TSCTで対側の微小陰影をフォローしていた例）
SQ, stage IIIA, Doubling time: 65日

Feng Li, MD, PhD
Shusuke Sone, MD
Hiroyuki Abe, MD, PhD
Heber MacMahon, MD
Kunio Doi, PhD

Index terms:
Cancer screening
Computed tomography (CT), thin-section
Lung neoplasms, CT, 60.1211
Lung neoplasms, diagnosis, 60.31, 60.32

Published online before print 10.1148/radiol.2333031018
Radiology 2004; 233:793–798

Abbreviations:
GGO = ground-glass opacity
PPV = positive predictive value

[1] From the Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, University of Chicago, 5841 S Maryland Ave, Chicago, IL 60637 (F.L., H.A., H.M., K.D.), and J. A. Azumi General Hospital, Ikeda, Nagano, Japan (S.S.). From the 2002 RSNA scientific assembly. Received June 30, 2003; revision requested September 9; final revision received February 27, 2004; accepted April 12. Supported in part by USPHS grant CA62625. Address correspondence to F.L. (e-mail: *fli@kurt.bsd.uchicago.edu*).

H.M. and K.D. are shareholders of R2 Technology, Los Altos, Calif. K.D is a shareholder of Deus Technology, Rockville, Md.

Author contributions:
Guarantors of integrity of entire study, F.L., S.S., K.D.; study concepts and design, F.L., K.D.; literature research, H.M., K.D.; clinical studies, F.L., S.S., H.A., H.M.; data acquisition, F.L., S.S.; data analysis/interpretation, F.L., H.A., H.M., K.D.; statistical analysis, F.L.; manuscript preparation, F.L.; manuscript definition of intellectual content, F.L., H.M., K.D.; manuscript editing, K.D., H.M.; manuscript revision/review and final version approval, all authors



# Malignant versus Benign Nodules at CT Screening for Lung Cancer: Comparison of Thin-Section CT Findings[1]

**PURPOSE:** To evaluate thin-section computed tomographic (CT) characteristics of malignant nodules on the basis of overall appearance (pure ground-glass opacity [GGO], mixed GGO, or solid opacity) in comparison with the appearance of benign nodules.

**MATERIALS AND METHODS:** Institutional review board approval and patient consent were obtained. Follow-up diagnostic CT was performed in 747 suspicious pulmonary nodules detected at low-dose CT screening (17 892 examinations). Of 747 nodules, 222 were evaluated at thin-section CT (1-mm collimation), which included 59 cancers and 163 benign nodules (3–20 mm). Thin-section CT findings of malignant versus benign nodules with pure GGO (17 vs 12 lesions), mixed GGO (27 vs 29 lesions), or solid opacity (15 vs 122 lesions) were analyzed. Fisher exact test for independence was used to compare differences in shape, margin, and internal features between benign and malignant nodules. Positive predictive value (PPV) was analyzed when a category was significantly different from the others.

**RESULTS:** Among nodules with pure GGO, a round shape was found more frequently in malignant lesions (11 of 17, 65%) than in benign lesions (two of 12, 17%; $P = .02$; PPV, 85%); mixed GGO, a subtype with GGO in the periphery and a high-attenuation zone in the center, was seen much more often in malignant lesions (11 of 27, 41%) than in benign lesions (two of 29, 7%, $P = .004$; PPV, 85%). Among solid nodules, a polygonal shape or a smooth or somewhat smooth margin was present less frequently in malignant than in benign lesions (polygonal shape: 7% vs 38%, $P = .02$; smooth or somewhat smooth margin: 0% vs 63%, $P < .001$), and 98% (46 of 47) of polygonal nodules and 100% (77 of 77) of nodules with a smooth or somewhat smooth margin were benign.

**CONCLUSION:** Recognition of certain characteristics at thin-section CT can be helpful in differentiating small malignant nodules from benign nodules.


Computed tomographic (CT) screening has increased the detection rate of early peripheral lung cancer in the United States and Japan (1,2). The results of the Early Lung Cancer Action Project, or ELCAP (1), suggested that nodules with pure (nonsolid) or mixed (partially solid) ground-glass opacity (GGO) at thin-section CT are more likely to be malignant than are those with solid opacity; among 44 nodules with GGO (19% of 233 nodules identified at baseline screening), 15 (34%) were confirmed to be malignant. On the other hand, most of the benign lesions were solid at CT, although some (approximately 15%) contained elements of GGO. According to the ELCAP data, 18% of nodules (five of 28) with pure GGO were malignant and 63% of nodules (10 of 16) with mixed GGO were malignant (1). To our knowledge, there are no previous studies that specifically compare thin-section CT characteristics between malignant lesions and benign lesions with pure GGO, mixed GGO, and solid opacity.

A 3-year lung cancer screening program has recently been completed in Japan by using low-dose CT and follow-up thin-section CT. We have previously reported that among 59

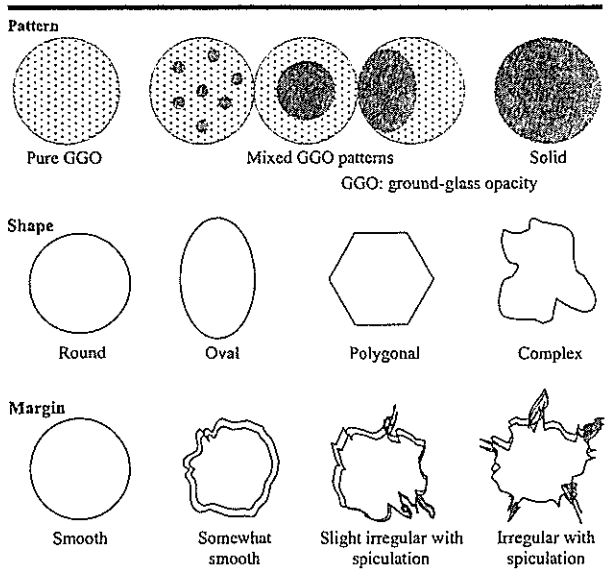



Figure 1. Typical appearance of the three patterns, four shapes, and four margins used to classify lesions in this study.

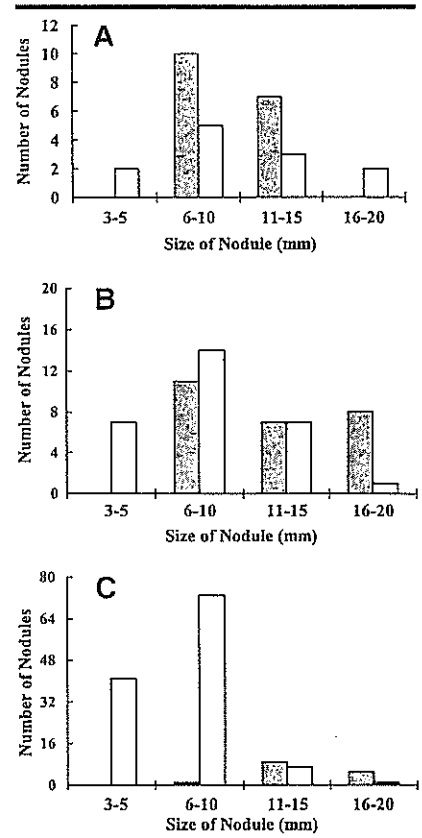



Figure 2. Graphs show distribution of sizes among, *A*, 29 nodules with pure GGO (17 malignant and 12 benign); *B*, 56 nodules with mixed GGO (27 malignant and 29 benign); and *C*, 137 nodules with solid opacity (15 malignant and 122 benign). Gray bars = malignant nodules, white bars = benign nodules. For pure and mixed GGO lesions, the size of benign nodules extensively overlaps that of malignant nodules in the 6–15-mm range.

small (6–20 mm) lung adenocarcinomas, only 16 nodules (27%) showed solid opacity and the rest (73%) showed pure or mixed GGO at thin-section CT in this screening program (2). In another study (3), thin-section CT characteristics were compared between 25 very small (≤10 mm) cancers, 24 of which were adenocarcinomas, and 40 benign lesions, most of which were solid nodules. We found that by using a single CT feature, namely polygonal shape, and a three-dimensional ratio (maximum transverse diameter to maximum z-axis dimension of a lesion, which was measured as the difference between the cephalic extent and the caudal extent of the lesion in coronal reformation) greater than 1.78, 100% specificity was shown for benign nodules (3). However, these features were not necessarily applicable to benign lesions with GGO, especially not to those larger than 10 mm. Thus, the purpose of our study was to evaluate the thin-section CT characteristics of malignant nodules on the basis of the overall appearance (pure GGO, mixed GGO, or solid opacity) compared with the appearance of benign nodules.

## MATERIALS AND METHODS

### Study Nodules

From May 1996 to March 1999, 17 892 examinations were performed in 7847 individuals (4288 men and 3559 women; mean age, 61 years) as part of an annual low-dose CT screening program for lung cancer in Nagano, Japan. A mobile unit equipped with a CT scanner (W950SR; Hitachi, Tokyo, Japan) was used to scan the chest with a tube current of 25 or 50 mA, a scanning time of 2 seconds per rotation of the x-ray tube (tube rotation time, 2 seconds), a table speed of 10 mm/sec (pitch of 2), 10-mm collimation, and a 10-mm reconstruction interval. The program was sponsored and supported by the Telecommunications Advancement Organization of Japan and was completed after 3 years. All subjects gave informed consent. Approval for review and research of the cases used in this study was obtained from our institutional review board at the University of Chicago.

Among those undergoing the examinations, 605 patients with 747 suspicious pulmonary nodules detected at low-dose CT underwent follow-up diagnostic CT. Diagnostic work-up CT, which included thin-section CT, was performed within 3 months of low-dose CT screening; follow-up CT examinations were performed at 3, 6, 12, 18, and 24 months, as needed. Most of the follow-up CT examinations were performed at Shinshu University Hospital, and some were performed at local hospitals. The results for follow-up work were accrued until December 1999.

The follow-up results for the 747 nodules include six categories, as follows: 76 primary lung cancers confirmed at biopsy; 11 atypical adenomatous hyperplasias confirmed at biopsy; 444 lesions, which included 167 resolved nodules, 230 nodules that were stable for 2 years or more, 38 nodules with benign-pattern calcifications (diffuse, central, popcorn, and laminar or concentric calcification), and nine nodules that were resected and confirmed as benign; 27 nodules with findings suspicious for malignancy at thin-section CT but not confirmed at biopsy; 176 nodules suspected of being benign but with insufficient follow-up; and 13 indeterminate nodules.

For this study, we used a database of thin-section CT images obtained from Shinshu University Hospital as part of the Nagano CT screening program for lung cancer. A helical scanner (HiSpeed Advantage; GE Medical Systems, Milwaukee, Wis) was used for scanning the nodules with a 200-mA tube current, 1 second per tube rotation, table speed of 1

TABLE 1
Thin-Section CT Findings in Malignant versus Benign Lesions with Pure GGO

| Feature | Malignant ($n$ = 17) | Benign ($n$ = 12) |
|---|---|---|
| Shape | | |
| Round | 11 | 2 |
| Oval | 3 | 1 |
| Polygonal | 0 | 3 |
| Complex | 3 | 6 |
| Margin | | |
| Smooth | 1 | 0 |
| Somewhat smooth | 9 | 5 |
| Slightly irregular with spiculation | 7 | 7 |
| Irregular with spiculation | 0 | 0 |

TABLE 2
Thin-Section CT Findings in Malignant versus Benign Lesions with Mixed GGO

| Feature | Malignant ($n$ = 27) | Benign ($n$ = 29) |
|---|---|---|
| Central opacity | | |
| Present | 11 | 2 |
| Absent | 16 | 27 |
| Air component | | |
| Present | 16 | 9 |
| Absent | 11 | 20 |
| Shape | | |
| Round | 10 | 2 |
| Oval | 1 | 2 |
| Polygonal | 3 | 9 |
| Complex | 13 | 16 |
| Margin | | |
| Smooth | 0 | 0 |
| Somewhat smooth | 4 | 7 |
| Slightly irregular with spiculation | 9 | 14 |
| Irregular with spiculation | 14 | 8 |

TABLE 3
Thin-Section CT Findings in Malignant versus Benign Solid Nodules

| Feature | Malignant ($n$ = 15) | Benign ($n$ = 122) |
|---|---|---|
| Air component | | |
| Present | 7 | 5 |
| Absent | 8 | 117 |
| Shape | | |
| Round | 7 | 39 |
| Oval | 0 | 24 |
| Polygonal | 1 | 46 |
| Complex | 7 | 13 |
| Margin | | |
| Smooth | 0 | 27 |
| Somewhat smooth | 0 | 50 |
| Slightly irregular with spiculation | 7 | 37 |
| Irregular with spiculation | 8 | 8 |

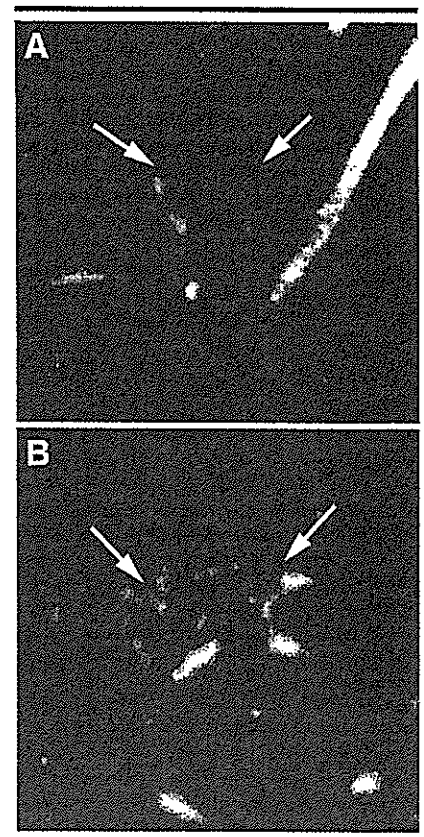



Figure 3. Transverse thin-section CT images. *A*, Image shows a malignant pure GGO lesion (adenocarcinoma) with a round shape (arrows). *B*, Image shows a benign pure GGO lesion (resolved within 3 months) with a polygonal-complex shape (arrows) that is confined to a secondary lobule.

mm/sec, 1-mm collimation, and 0.5-mm interval with a bone reconstruction algorithm. The database consisted of studies performed in 222 patients with 222 confirmed malignant or confirmed benign nodules, which were small in size (3–20 mm) on the first thin-section CT image obtained within 3 months of low-dose CT screening. Among the 222 patients, there were 14 patients with two nodules in different lung lobes, in which case the larger of the two nodules was selected for this study. Patients with two nodules in the same lung lobe and patients with more than two nodules were not included. On thin-section CT images, nonnodular lesions such as linear or scattered opacities, which had been regarded as suspicious on the original 10-mm collimation screening CT images, were excluded from the analysis. Nodules with benign-pattern calcifications were also excluded. This database contained cases of 96 pulmonary nodules that were used in two previous studies (2,3).

Among the 222 patients (mean age, 62.4 years; age range, 30–84 years), there were 119 men (mean age, 62.8 years; age range, 30–84 years) and 103 women (mean age, 61.9 years; age range, 34–75 years).

## Data Analysis

Thin-section CT images for the 222 nodules were displayed and interpreted with use of "stacked" mode on a monochrome cathode ray tube monitor at a width and level of 1500 HU and −550 HU, respectively. The images of 222 nodules were randomly arranged for a reading sequence, and the final diagnosis for the nodules, which included the his-

topathologic results, was blinded to the radiologists. Three radiologists with 20, 18, and 17 years of experience in general radiology (F.L. and H.A. included) independently viewed these images and subjectively classified the nodules as one of three patterns: pure GGO, mixed GGO, or solid opacity. They also independently determined the overall shape (round, oval, polygonal, or complex) and margin (smooth, somewhat smooth, somewhat irregular with slight spiculation, or irregular with spiculation) of the nodules, as well as the internal features. Internal features included a specific mixed GGO pattern characterized by GGO in the periphery, with a high-attenuation zone in the center and the presence or absence of air (air bronchogram, cavitation, or focal emphysema) within the nodule on thin-section CT images. The typical appearance of the three patterns, four shapes, and four margins used to classify the lesions is illustrated in Figure 1.

For pattern, shape, and margins of the nodules, the same judgment was made by all three radiologists for 75%, 40%, and 13% of cases, respectively, and the same judgment was made by any two of the radiologists for 99%, 91%, and 73% of cases, respectively. Two radiologists (F.L., H.A.) worked together to reach a consensus for the remaining 83 features in 76 nodules; these nodules were initially classified differently by each of the three radiologists. For internal features, the same judgment was made by all three radiologists in 67% of cases and by any two radiologists in 100% of cases. The final decision regarding the CT findings was based on the consensus of at least two radiologists. The mean size (average length and width) and clinical outcome for 222 nodules were recorded by one radiologist (F.L.).

### Statistical Analysis

Statistical analysis was performed by using the Student *t* test for comparison of differences in size between benign and malignant nodules. The $\chi^2$ test for independence was used independently for comparison of the differences in patterns (nodules with and those without GGO) between the benign nodules and the malignant nodules. The data presented in Tables 1–3 were analyzed first by using the Fisher exact test for independence to determine whether there were any significant differences in the proportion of malignant lesions and benign lesions in the categories of shape, margin, and internal features. If such differences were established (the difference was significant at $P \leq .05$), additional Fisher exact tests were performed to determine which categories were significantly different from the others. Fisher exact test was used instead of $\chi^2$ test because of the small sample size. Positive predictive value (PPV) was further analyzed when a category was significantly different from the others.

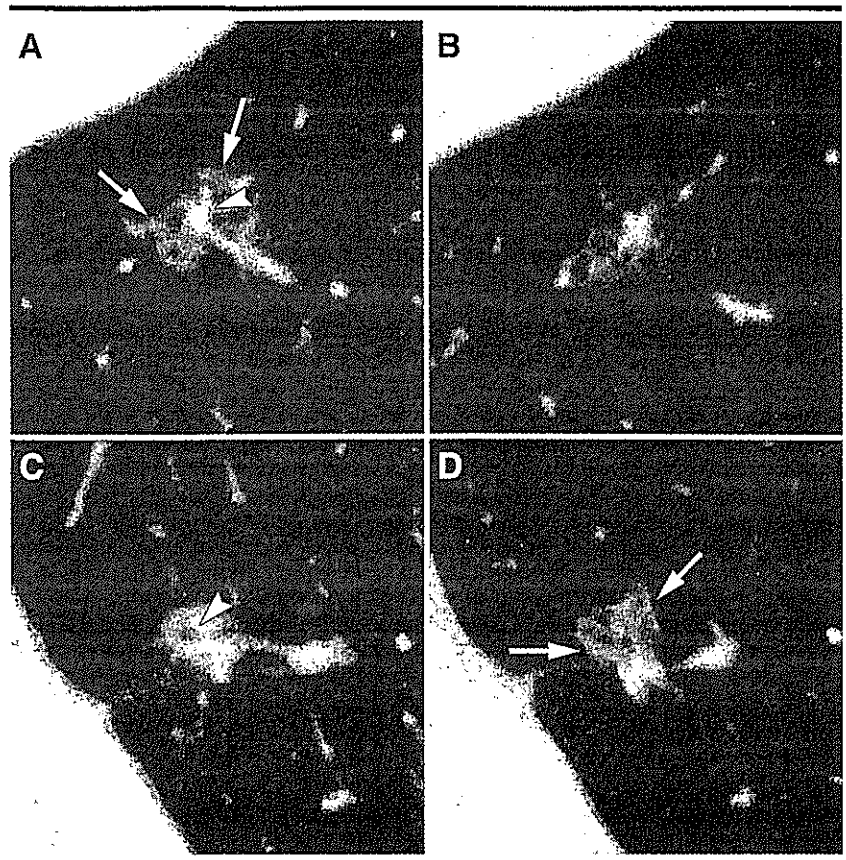



Figure 4. Transverse thin-section CT images. *A*, *B*, Images show a malignant mixed GGO lesion (adenocarcinoma) with irregular margins. The nodule shows both GGO in the periphery (arrows) and a high-attenuation zone (arrowhead) in the center. *C*, *D*, Images show a benign mixed GGO lesion (nodular fibrosis) with irregular margins. In *C*, a small air collection (arrowhead) is seen in the nodule. In *D*, the nodule (arrows) is seen on another section.

## RESULTS

Of the 222 patients evaluated, 59 (27 men and 32 women; mean age, 64.6 years) had malignant nodules and 163 (92 men and 71 women; mean age, 61.6 years) had benign nodules. The mean size of the 59 malignant nodules (12.3 mm) was larger than that of the 163 benign nodules (7.2 mm, $P < .001$). Among 59 malignant nodules, there were 17 with pure GGO, 27 with mixed GGO, and 15 with solid opacity. Among 163 benign nodules, 12 showed pure GGO, 29 showed mixed GGO, and 122 showed solid opacity. The number of lesions with GGO was greater in the group of malignant nodules than in the group of benign nodules ($P < .001$).

All 17 malignant nodules with pure GGO were well-differentiated adenocarcinomas. Among 27 malignant nodules with mixed GGO, 26 were well-differentiated adenocarcinomas and one was a moderately differentiated adenocarcinoma. Of the 15 malignant nodules with solid opacity, four were well-differentiated adenocarcinomas, seven were other adenocarcinomas, two were squamous cell carcinomas, and two were small cell carcinomas. All 12 benign nodules with pure GGO had resolved at the 3-month follow-up examination. Among 29 benign nodules with mixed GGO, nodular fibrosis was confirmed at surgery in three cases, was re-

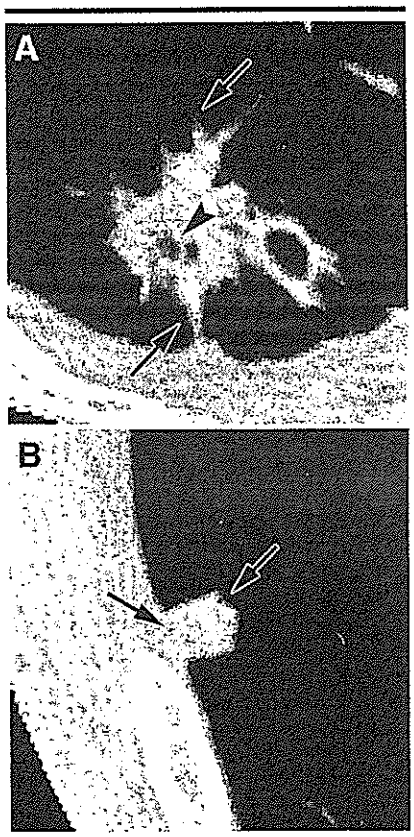



Figure 5. Transverse thin-section CT images. *A*, Image shows a malignant nodule (squamous cell carcinoma) with air components (arrowhead) and an irregular margin and gross spiculation (arrows). *B*, Image shows a small benign solid nodule (stable for more than 2 years) with a polygonal shape (arrows) and somewhat smooth margin.

solved at 3 months or more of follow-up in 17 cases, and showed no change for 2 years or more in nine cases. Among the 122 benign solid nodules, five cases (one case each of inflammatory granuloma, cryptococcoma, focal organizing pneumonia, inflammatory pseudotumor, and sclerosing hemangioma) were confirmed at surgery, 19 cases were resolved at 3 months or more of follow-up, and 98 cases showed no change for 2 years or more. All malignant nodules were confirmed at surgery.

The distribution of sizes among 29 nodules with pure GGO, 56 with mixed GGO, and 137 with solid opacity is shown in Figure 2. For GGO lesions, there was extensive overlap between the size of benign nodules and that of malignant nodules. On the other hand, for solid lesions, there was relatively limited overlap between the size of benign nodules and that of malignant nodules.

Table 1 shows the thin-section CT findings for malignant versus benign lesions with pure GGO; Figure 3 shows a malignant nodule and a benign nodule with pure GGO obtained at thin-section CT. The overall Fisher exact test indicated a significant association between lesion shape and malignancy ($P = .008$) but indicated no significant association between margins and malignancy ($P = .826$). At further examination, we found a significant association between malignancy and round nodules ($P = .022$); the number of round nodules was greater in the malignant group (65%, 11 of 17) than in the benign group (17%, two of 12) of pure GGO lesions. If round shape was used to discriminate between malignant lesions and benign lesions with pure GGO, the PPV (probability that a nodule is malignant, given that it is round) of such a test would be 85% (95% confidence interval: 54.55%, 98.08%) in this data set.

Table 2 lists thin-section CT findings for malignant lesions versus benign lesions with mixed GGO; Figure 4 shows a malignant nodule and a benign nodule with mixed GGO obtained at thin-section CT. The overall Fisher exact test again showed a significant association between nodule shape and malignancy ($P = .020$) but showed no significant association between margins and malignancy ($P = .174$). The association between round nodules and malignancy was found to be significant ($P = .009$), and the proportion of round nodules was higher among malignant lesions (37%, 10 of 27) than among benign lesions (7%, two of 29). The PPV was 83% (95% confidence interval: 51.59%, 97.91%). Furthermore, the presence of central opacity with mixed GGO was significantly associated with malignancy ($P = .004$), with a higher proportion of nodules with this feature in the malignant group (41%, 11 of 27) than in the benign group (7%, two of 29). The PPV of this test was 85% (95% confidence interval: 54.55%, 98.08%). However, the presence of air components within lesions was not significantly associated with malignancy ($P = .059$).

Table 3 lists thin-section CT findings for malignant versus benign solid nodules; Figure 5 shows malignant nodules and benign nodules obtained at thin-section CT. Fisher exact test showed a significant association between shape and malignancy ($P < .001$), as well as between margins and malignancy ($P < .001$). However, a round shape was not found to be associated with malignancy in solid nodules ($P = .262$), which is in contrast to the results found with pure and mixed GGO lesions. An oval shape was not significantly associated with malignancy ($P = .073$). The association between a complex shape and malignancy was found to be significant ($P = .002$)—the proportion of nodules with complex shape was higher among malignant lesions (47%, seven of 15) than among benign lesions (11%, 13 of 122). However, the PPV of this test was only 35% (95% confidence interval: 15.39%, 59.22%). The proportion of nodules with a polygonal shape was greater among benign lesions (38%, 46 of 122) than among malignant lesions (7%, one of 15; $P = .019$). There were 47 polygonal nodules, 46 (98%) of which were benign. When the margin classifications were dichotomized into "smooth or somewhat smooth" and "slightly irregular or irregular" categories, there was a significant difference between benign nodules and malignant nodules ($P < .001$). The proportion of smooth or somewhat smooth margins among malignant lesions was lower (0%, none of 15) than it was among benign lesions (63%, 77 of 122). There were 77 smooth or somewhat smooth nodules, and all 77 were benign. Furthermore, the presence of air components within these solid lesions was significantly associated with the malignant group (47%, seven of 15; $P < .001$) in comparison with the benign group (4%, five of 122). The PPV of this test was 58% (95% confidence interval: 27.67%, 84.83%).

## DISCUSSION

Comparison of various CT features such as contour, margins, and internal characteristics of pulmonary nodules with pathologic specimens can be helpful for developing criteria to distinguish between cancers and benign lesions (1,2,4–7). In CT screening programs, however, most benign nodules are not confirmed at pathologic diagnosis. Because of this limitation, we were not able to make a detailed radiologic-pathologic comparison. Therefore, we chose to investigate two internal patterns, namely *(a)* nodules with both GGO in the periphery and a high-attenuation zone in the center and *(b)* nodules with an area of air, such as an air bronchogram, that is frequently found in small well-differentiated adenocarcinomas (2,5). Also, we classified all nodules into one of four subcategories of shape and margins on the basis of the predominant CT appearance. In our

study, we found that differences in the CT features between benign lesions and malignant lesions were observed for each of the three patterns on thin-section CT images.

Results of previous clinical CT studies (8–10) have shown that malignant nodules commonly contain solid opacity and that benign nodules have higher attenuation, often with visible calcifications, than do malignant nodules. Siegelman et al (10) reported that 61% of 279 benign nodules (including 153 nodules with diffuse calcifications) had smooth or moderately smooth margins and 65% of 283 primary malignant tumors had irregular shapes with spiculation. Kuriyama et al (5), in a study of 20 peripheral lung cancers and 20 benign nodules less than 20 mm in diameter, reported that an air bronchogram was not observed as frequently in small benign lesions, such as hamartoma and tuberculoma, as it was in adenocarcinomas.

The number of solid benign nodules was much greater than the total number of malignant nodules in our database, which was obtained from a lung cancer CT screening program, and the frequency of some features, such as internal air bronchograms, a complex shape, and an irregular margin, was much less in common in benign lesions than in malignant lesions. However, these observations do not necessarily mean that these features are reliable for differentiating benign nodules from malignant nodules, because the absolute numbers of benign nodules with such features may be comparable to the numbers of malignant nodules with similar features. For example, the frequency of an irregular margin in solid nodules was 7% (eight of 122) for benign nodules and 53% (eight of 15) for malignant nodules. However, if a radiologist encountered such a case in a screening examination, there would be an approximately 50% (eight of 16) likelihood that the lesion was malignant, if all other factors were equal. We found that a polygonal shape or a smooth or somewhat smooth margin (98%–100% likelihood of benignity) could be more helpful for differentiating solid benign nodules from malignant nodules than would internal air bronchograms, a complex shape, or an irregular margin.

There were some limitations to this study. For instance, no malignant lesions 5 mm or smaller were found; this is probably because the database used here was compiled from images obtained with low-dose single-detector row CT at a 10-mm section thickness. Second, many of the benign GGO lesions detected at the initial screening CT had resolved before thin-section diagnostic CT was performed. In a previous study, we reported that among 108 benign nodules (54, 27, and 27 of which showed pure GGO, mixed GGO, and solid opacity, respectively, at low-dose CT), 92 (85%) resolved within 3 months (11). Also, a large variance was noted in the judgment for CT features by three radiologists, especially for margins of the nodules; this is probably because most nodules used in current study were smaller than 10 mm.

The margins and size of nodules were not useful for differentiating benign from malignant GGO lesions in this series, and benign lesions with GGO were more difficult to distinguish from malignant nodules than were those with solid opacity. However, certain features, such as a round shape or a combination of GGO in the periphery with a high-attenuation zone in the center, were observed much more frequently in malignant GGO nodules. Therefore, we believe that familiarity with the different features of benign nodules and malignant nodules can be useful to radiologists in the management of indeterminate nodules. Also, short-term follow-up imaging can be helpful for differentiating benign from malignant nodules with GGO patterns, because all 12 of the benign pure GGO lesions in this series, as well as the majority of benign mixed GGO lesions, had partially or completely resolved within 3 months.

**Acknowledgments:** The authors are grateful to Chaotong Zhang, MD, for participating as an image reviewer, Qiang Li, PhD, for helpful suggestions, Masha Kocherginsky, PhD, for assistance with statistical analysis, Roger Engelmann, MS, for his useful work on the display software, and Elisabeth Lanzl for the English editing of the manuscript.

**References**

1. Henschke CI, Yankelevitz DF, Mirtcheva R, et al. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. AJR Am J Roentgenol 2002; 178:1053–1057.
2. Yang ZG, Sone S, Takashima T, et al. High-resolution CT analysis of small peripheral lung adenocarcinomas revealed on screening helical CT. AJR Am J Roentgenol 2001; 176:1399–1407.
3. Takashima S, Sone S, Li F, et al. Small solitary pulmonary nodules (< or = 1 cm) redetected at population-based CT screening for lung cancer: reliable high-resolution CT features of benign lesions. AJR Am J Roentgenol 2003; 180:955–964.
4. Zwirewich CV, Vedal S, Miller RR, Müller NL. Solitary pulmonary nodule: high-resolution CT and radiologic-pathologic correlation. Radiology 1991; 179:469–476.
5. Kuriyama K, Tateishi R, Doi O, et al. Prevalence of air bronchograms in small peripheral carcinomas of the lung on thin-section CT. AJR Am J Roentgenol 1991; 156:921–924.
6. Li F, Sone S, Takashima S, et al. Correlations between high-resolution computed tomographic, magnetic resonance and pathological findings in cases with non-cancers but suspicious lung nodules. Eur Radiol 2000; 10:1782–1789.
7. Kohno N, Ikezoe J, Johkoh T, et al. Focal organizing pneumonia: CT appearance. Radiology 1993; 189:119–123.
8. Proto AV, Thomas SR. Pulmonary nodules studied by computed tomography. Radiology 1985; 156:149–153.
9. Zerhouni EA, Stitik FP, Siegelman SS, et al. CT of the pulmonary nodule: a cooperative study. Radiology 1986; 160:319–327.
10. Siegelman SS, Khouri NF, Leo FP, Fishman EK, Braverman RM, Zerhouni EA. Solitary pulmonary nodules: CT assessment. Radiology 1986; 160:307–312.
11. Li F, Sone S, Takashima S, Maruyama Y, Hasegawa M, Yang ZG. Roentgenologic analysis of 108 non-cancerous focal lung lesions detected in screening CT for lung cancer. Jpn J Lung Cancer 1999; 39:369–380.

# Radiologists' Performance for Differentiating Benign from Malignant Lung Nodules on High-Resolution CT Using Computer-Estimated Likelihood of Malignancy

Feng Li[1]
Masahito Aoyama[2]
Junji Shiraishi[1]
Hiroyuki Abe[1]
Qiang Li[1]
Kenji Suzuki[1]
Roger Engelmann[1]
Shusuke Sone[3]
Heber MacMahon[1]
Kunio Doi[1]

**OBJECTIVE.** The purpose of our study was to evaluate whether a computer-aided diagnosis (CAD) scheme can assist radiologists in distinguishing small benign from malignant lung nodules on high-resolution CT (HRCT).

**MATERIALS AND METHODS.** We developed an automated computerized scheme for determining the likelihood of malignancy of lung nodules on multiple HRCT slices; the likelihood estimate was obtained from various objective features of the nodules using linear discriminant analysis. The data set used in this observer study consisted of 28 primary lung cancers (6–20 mm) and 28 benign nodules. Cancer cases included nodules with pure ground-glass opacity, mixed ground-glass opacity, and solid opacity. Benign nodules were selected by matching their size and pattern to the malignant nodules. Consecutive region-of-interest images for each nodule on HRCT were displayed for interpretation in stacked mode on a cathode ray tube monitor. The images were presented to 16 radiologists—first without and then with the computer output—who were asked to indicate their confidence level regarding the malignancy of a nodule. Performance was evaluated by receiver operating characteristic (ROC) analysis.

**RESULTS.** The area under the ROC curve ($A_z$ value) of the CAD scheme alone was 0.831 for distinguishing benign from malignant nodules. The average $A_z$ value for radiologists was improved with the aid of the CAD scheme from 0.785 to 0.853 by a statistically significant level ($p = 0.016$). The radiologists' diagnostic performance with the CAD scheme was more accurate than that of the CAD scheme alone ($p < 0.05$) and also that of radiologists alone.

**CONCLUSION.** CAD has the potential to improve radiologists' diagnostic accuracy in distinguishing small benign nodules from malignant ones on HRCT.

CT screening has led to early detection of peripheral lung cancer and also detection of a large number of false-positives (i.e., noncalcified benign nodules) [1–5]. The false-positive rate at screening has been reported as 87–93% with low-dose single-detector CT at 10-mm slice thickness [1–3] and 98–99% with single-detector CT or MDCT at 5-mm slice thickness [4, 5]. Also, simultaneous or additional diagnostic high-resolution CT (HRCT) is needed for the distinction between malignant and benign lung nodules detected as suspicious or indeterminate lesions on screening CT [1–5]. This high false-positive rate because of benign nodules is likely to reduce the benefit of CT screening for early detection of lung cancer [6]. Therefore, it is important to differentiate benign from malignant nodules to reduce the number of false-positive findings on screening CT and to reduce follow-up examinations for diagnostic HRCT.

We developed an automated computerized scheme [7] for determination of the likelihood measure of malignancy by using various objective features of the nodules in our a database of thick-section low-dose CT; one or two slices were used for image analysis of each nodule. The low-dose CT database consisted of 489 nodules obtained from a mass screening for lung cancer in Nagano, Japan [2]. All of these nodules were considered as suspicious or indeterminate lesions when detected by radiologists on low-dose CT screening. With the use of receiver operating characteristic (ROC) analysis, our computerized scheme achieved an area under the ROC curve ($A_z$ value) of 0.846 for distinction between 76 malignant and 413 benign lung nodules.

Recently, we further developed another computerized scheme for distinction be-

Received November 10, 2003; accepted after revision January 30, 2004.

Supported in part 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.

[1]Department of Radiology, Kurt Rossmann Laboratories for Radiologic Image Research, MC-2026, The University of Chicago, 5841 S Maryland Ave., Chicago, IL 60637. Address correspondence to F. Li (fli@kurt.bsd.uchicago.edu).

[2]Department of Intelligent Systems, Faculty of Information Sciences, Hiroshima City University, Hiroshima 731-3194, Japan.

[3]Azumi General Hospital, Ikeda, Nagano 399-8695, Japan.

*AJR* 2004;183:1209–1215

0361–803X/04/1835–1209

tween malignant and benign lesions derived from multiple slices of HRCT (1-mm collimation) based on 2D and 3D volume data. The HRCT database consisted of 244 small noncalcified (3–20 mm) nodules obtained as part of follow-up diagnostic work for suspicious or indeterminate lesions detected on low-dose CT in the same screening program.

In the present study, we assessed observer performance using ROC analysis to evaluate the effectiveness of our computer-aided diagnosis (CAD) scheme to assist radiologists in distinguishing small benign from malignant lung nodules in various patterns at HRCT. The malignant lung cancers included nodules with pure ground-glass opacity, mixed ground-glass opacity, and solid opacity; the benign nodules were selected by matching their size and pattern to the cancers on HRCT in this observer study.

## Materials and Methods

Our institutional review board approved the use of this database and the participation of radiologists in this observer performance study. Informed consent for use of cases was waived. Informed consent for the observer performance study was obtained from all observers.

### *Database*

The diagnostic HRCT database used in this study consisted of 59 patients (27 men, 32 women; mean age, 64.6 years) with 61 malignant nodules and 169 patients (99 men, 70 women, mean age 61.6 years) with 183 benign nodules. The database was obtained as part of an annual 3-year low-dose CT screening for lung cancer from 17,892 examinations on 7,847 individuals in Nagano, Japan [2]. HRCT scans were obtained on a helical scanner (HiSpeed Advantage, GE Healthcare) with a standard tube current (200 mA) to cover the entire nodule lesion, 1-mm collimation, and a bone reconstruction algorithm with a 0.5-mm interval.

Two features concerning the size and pattern type of the pulmonary nodules on HRCT were subjectively determined by radiologists for the purpose of grouping nodules in our database. The mean size (average length and width) was recorded by one radiologist. The three types of patterns of these nodules—pure ground-glass opacity, mixed ground-glass opacity, and solid opacity—were viewed independently and grouped by three radiologists without knowledge of the final diagnosis, and a consensus was reached through discussion. Nodules with benign-pattern calcifications (diffuse, central, popcorn, and laminar, or concentric calcification) were excluded. The range of nodule sizes for the 61 malignant and 183 benign nodules was 6–19 mm (mean, 12 mm) and 3–20 mm (mean, 7 mm), respectively. Among the 61 malignant nodules, there were 18 nodules with pure ground-glass opacity, 28 with mixed ground-glass opacity, and 15 with solid opacity, whereas 183 benign nodules included 12 with pure ground-glass opacity, 30 with mixed ground-glass opacity, and 141 with solid opacity.

All malignant nodules were primary lung cancers confirmed by surgery, including 49 well-differentiated adenocarcinomas, eight other adenocarcinomas, two squamous cell carcinomas, and two localized small cell carcinomas. Among the 183 benign nodules, nine (four cases of nodular fibrosis; and one case each of inflammatory granuloma, cryptococcoma, focal organizing pneumonia, inflammatory pseudotumor, and sclerosing hemangioma) were confirmed by surgery, 51 had resolved at follow-up examination, and 123 showed no change for 2 or more years.

### *CAD*

With our CAD scheme, the nodules were segmented automatically using a dynamic programming technique [7]. Forty-one and 15 image features based on 2D and 3D volume data, respectively, were determined from quantitative analysis of the nodule outline and pixel values. Linear discriminant analysis was used to distinguish benign from malignant nodules. The performance of this CAD scheme was evaluated on the basis of a leave-one-out testing method by use of 61 malignant and 183 benign lung nodules. For the input of the linear discriminant analysis, we selected many combinations from 56 features and two clinical parameters (patient age and sex). The following features were used in this study: effective diameter, contrast of the segmented nodule on the HRCT image, overlap measures of two gray-level histograms for the inside and outside regions of the segmented nodule on the HRCT image, overlap measures of two gray-level histograms for the inside and outside regions of the segmented nodule on the edge-gradient image, radial gradient index for the inside region of the segmented nodule on the HRCT image, peak value of the histogram for the inside region of the segmented nodule on the edge-gradient image, pixel value at the peak of the histogram for the inside region of the segmented nodule on the edge-gradient image, and pixel value at the peak of the histogram for the inside region of the segmented nodule on the HRCT image.

Our computerized classification method outputs a percentage, from 1% to 99%, that indicates the likelihood of malignancy. The performance of the classification scheme yielded an $A_z$ value of 0.937 (0.919 for nodules with pure ground-glass opacity, 0.852 for nodules with mixed ground-glass opacity, and 0.957 for solid nodules) for distinction between 61 malignant and 183 benign lung nodules.

### *Observer Study*

The data used in this observer study consisted of 28 malignant nodules that were randomly selected from the 61 primary lung cancers and 28 benign nodules that were selected from the 183 benign nodules by matching in size and pattern to the cancers. For both malignant and benign lesions, nine nodules ranged from 6 to 10 mm and 19 nodules ranged from 11 to 20 mm. The patterns involved were eight nodules with pure ground-glass opacity, 12 with mixed ground-glass opacity, and eight with solid opacity. Examples of cases used for this observer study are shown in Figure 1.

Sixteen radiologists participated in this observer study. The 16 radiologists, seven chest radiologists and nine other radiologists, have a mean of 14 years of experience (range, 7–26 years). Consecutive region-of-interest HRCT images for each nodule were displayed for interpretation using the cine mode on a cathode ray tube monitor (1,280 × 1,024 resolution). The window settings were initially at a width of 1,500 H and a level of −550 H, but the settings could be adjusted by the observer. In addition, zooming capability was provided. Two clinical parameters (patient age and sex) were displayed to the observer on the monitor.

The observers were told that the purpose of this observer study was to assist radiologists in distinguishing benign from malignant lesions on HRCT by using a CAD scheme. The instructions for the observers were an explanation of the role of CAD output as a second opinion. The observers were told that 28 malignant lesions (6–10 mm, nine cases; 11–20 mm, 19 cases; pure ground-glass opacity, eight cases; mixed ground-glass opacity, 12 cases; and solid opacity, eight cases) and 28 benign lesions (matched in size and pattern to the malignant lesions) were included in this study and that the sensitivity and specificity of our CAD scheme, using a threshold of 0.50 (50%) likelihood of malignancy, are 80% and 75%, respectively.

The observers were instructed to click on a bar (left, benignancy; right, malignancy) on the screen using a mouse to indicate confidence level regarding the malignancy (or benignancy) of a lesion first without and then with computer output, and after indicating your confidence (without and with CAD), click on one of the four following clinical actions: return to annual screening; follow-up in 6 months; follow-up in 3 months; or biopsy or surgery.

For a training session before the test, we provided five cases so that the observers could learn how to operate the cine mode interface and how to take into account the computer output in their decision. The review time was not limited. The average review time was 46 min (range, 28–100 min).

### *Data Analysis*

The confidence level ratings from each observer were analyzed using receiver operating characteristic (ROC) methodology, and a quasimaximum likelihood estimation of the binormal distribution was fitted to the radiologists' confidence ratings [8]. The statistical significance of the difference in $A_z$ values between observer interpretations without and with the CAD scheme was tested using the Dorfman-Berbaum-Metz method [9]; this method included both observer variation and case sample variation by means of an analysis-of-variance approach. The sta-

tistical significance of the difference in $A_z$ values between the computer outputs and observer interpretations (without and with the CAD scheme) was tested by means of confidence interval method by taking into account observer variation alone [10]. The effect of the computer output on the rating scores and also the change in scores that were due to the use of the CAD scheme were analyzed. The distributions of the radiologists' ratings and of the computer outputs were compared for the malignant and benign nodules.

The statistical significance of the difference in clinical actions between the beneficial and detrimental effect of the CAD scheme for each of the malignant and benign nodules was estimated using the Student's paired *t* test for 16 radiologists.

## Results

For the cases selected for this observer study, the $A_z$ value of the CAD scheme alone was 0.831 for distinguishing 28 malignant and 28 benign nodules (0.910 for nodules with pure ground-glass opacity, 0.814 for nodules with mixed ground-glass opacity, and 0.783

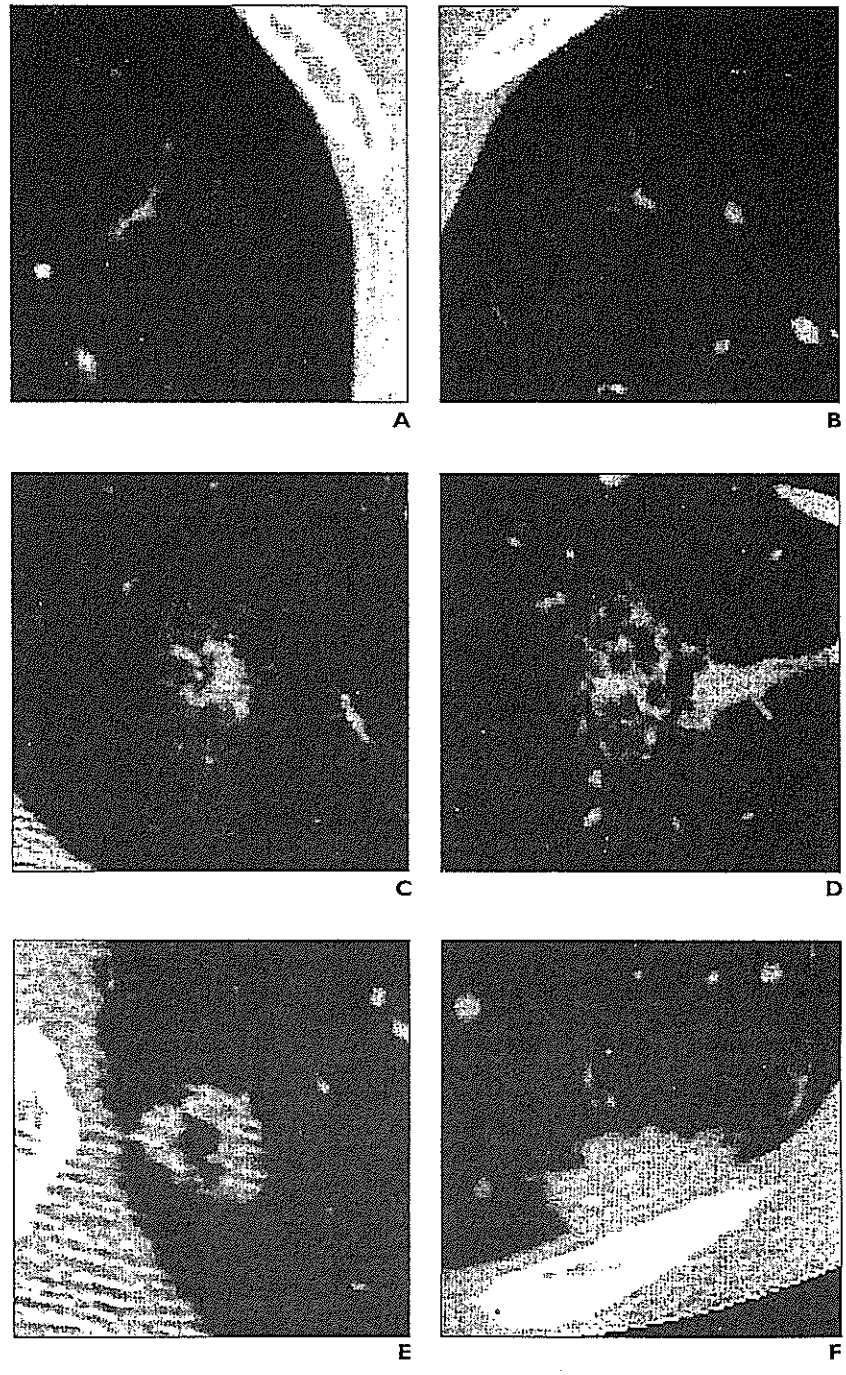



**Fig. 1.**—Radiologists' average ratings without and with computer output for six cases used in observer study. Note that difference in likelihood of malignancy between computer output and initial radiologists' ratings was not large in cases shown here. Radiologists' interpretation with computer-aided diagnosis (CAD) scheme was, in general, more accurate than radiologists without CAD scheme in most malignant and benign nodules.

**A,** High-resolution CT (HRCT) scan of 55-year-old woman with lung cancer shows pure ground-glass opacity. Computer output was 0.66; radiologists' ratings without CAD, 0.64; and radiologists' ratings with CAD, 0.71.

**B,** HRCT scan of 57-year-old woman with benign nodule shows pure ground-glass opacity. Computer output was 0.24; radiologists' ratings without CAD, 0.32; and radiologists' ratings with CAD, 0.27.

**C,** HRCT scan of 73-year-old man with lung cancer shows mixed ground-glass opacity. Computer output was 0.90; radiologists' ratings without CAD, 0.75; and radiologists' ratings with CAD, 0.85.

**D,** HRCT scan of 79-year-old man with benign nodule shows mixed ground-glass opacity. Computer output was 0.57; radiologists' ratings without CAD, 0.48; and radiologists' ratings with CAD, 0.56.

**E,** HRCT scan of 57-year-old man with lung cancer shows solid opacity. Computer output was 0.78; radiologists' ratings without CAD, 0.66; and radiologists' ratings with CAD, 0.76.

**F,** HRCT scan of 68-year-old man with benign nodule shows solid opacity. Computer output was 0.36; radiologists' ratings without CAD, 0.37; and radiologists' ratings with CAD, 0.36.