Confidence distribution of presence or absence of an abnormality judged by an interpreter (CTS5 & 6)

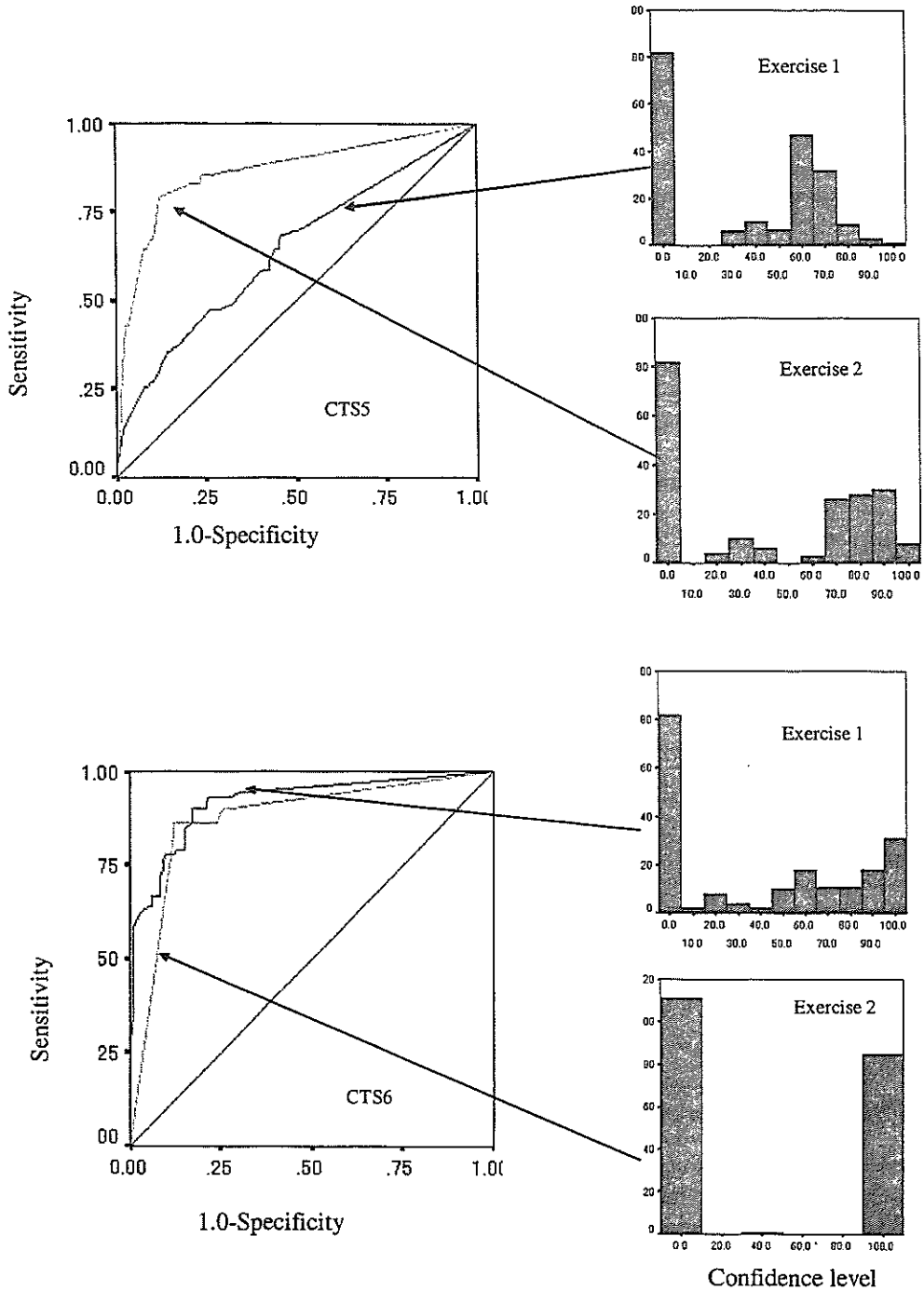

Figure.5 Influence of training on confidence distribution

in reality, the interpreter must choose correctly between the two opposing assessments produced by the screener and this type of CAD system. As shown in Figure 3, the fact that the Az for Exercise 3 increased overall in comparison with that for Exercise 2 suggests that many of the CTSs (95%) were able to choose the correct assessment effectively out of the two opposing assessments. On the other hand, the effectiveness of using the CAD system as a reference was also observed for the physician who participated in Exercise 1 and exhibited the highest diagnostic ability. This fact indicates that the performance or complementarity of the hypothetical CAD system was excellent. It must also be stressed that the way in which the hypothetical CAD system detected either no abnormality or a single abnormality, and that the result of the interpreters' diagnosis of the presence of an abnormality was also either zero or one, is not indicative of reality. In Exercise 3, when the results given by both the CAD system and the interpreters were simplified to zero or one, the issue for the interpreter was to compare his or her own opinion with the result of the CAD system and choose the correct one. In contrast, the general tendency among actual CAD systems currently under development is to produce comparatively more false positive results than those assessed by interpreters[11]. It will be necessary to determine the level of complementarity with interpreters offered by actual CAD systems. Research is also necessary on how to use a CAD system for reference in a situation where the number of true pathological changes that an interpreter should detect may not be limited to zero or one because two or more may be present and whether CTSs can use their own assessments to illuminate the results of the CAD system and select those that are correct to improve their diagnostic accuracy even when a CAD system indicates several potential abnormalities, including false positives[12]. If the hypothetical CAD system used as an experimental reference in Exercise 3 were to possess this type of performance, even the CTSs would be able to achieve a diagnostic level at least as good as the physicians actually employed in this task, and our results indicate this as a goal for the developers of CADs.

All interpretation by both CTSs and physicians was carried out indoors under conditions of low ambient lighting. The PG that was compared with the CTSs interpreted images displayed on a CRT while wearing an eye camera for the purpose of gathering data on their line of vision. The CTSs, on the other hand, interpreted images projected onto a screen by a PC projector. The observation conditions for the CTSs were variable in that some were too close to the screen, others too far away, and still others viewed the images from a direction not perpendicular to the screen. It is not possible to determine which viewing conditions were more effective, those for the CTSs or for the PG. Figure 3 compares the diagnostic accuracy of the CTSs and the PG, but as the interpreting conditions were different it is of limited use.

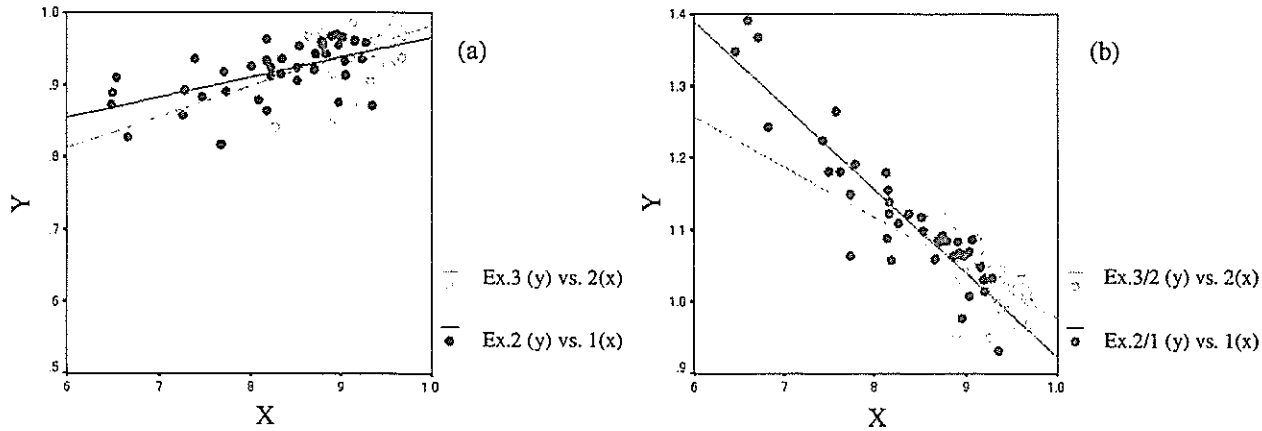

Fig.6 Effect on CT image interpretation of training and use of CAD

The physician who sat in with the CTSs during the exercises carried out the interpretation under the same conditions as they did. The Az achieved by this physician as a result was outstanding compared not only with those of the CTSs but also with those of the PG. It may be inferred that this reflects the result of the physician being placed in a difficult position, but with an interpreting ability which was directly compared with that of the CTSs, in addition to the existing abilities of the PG in interpreting. Immediately before Exercise 3, all interpreters were given a copy of Figure 4 showing their performance stating specificity and sensitivity, and were informed that if they combined their own assessment skillfully with the result given by the CAD system they could obtain the high scores shown in the top right of the diagram. Accordingly, the results of Exercise 3 reflect the effect of having been set a goal to achieve. We may also infer that these results reflect the effects of self-training as a result of the practice before Exercise 1, the lectures in CT image diagnostics received between Exercises 1 and 2, the training in interpreting methods immediately before Exercise 2, and

of interpreting the same cases (nos. 1–100) three times in succession, even though there was a lapse of one month between Exercises 1 and 2 and three months between Exercises 2 and 3. The result of a comparison between the scores obtained by the CTSs under these unusual conditions with those of the PG (Figure 3) cannot therefore be used immediately to generalize concerning the potential of CTSs as CT screeners. Furthermore, the CT images used for interpretation in this research consisted of one image slice for each one of a collection of 100 cases. Normally, a single case generates a high volume of around 30 CT images. In future it will be necessary to investigate the potential of screeners by using a CT image database that more closely approximates to reality.

Figure 5 shows a sample confidence distribution for the answers given by CTSs 5 and 6 during Exercises 1 and 2. The confidence levels of CTS 5 were continuously distributed in Exercise 1, but as a result of training the distribution shifted toward 0% or 100%, and the ROC curve improved. The confidence levels of CTS 6 also had a continuous distribution in Exercise 1, but most of the answers given in Exercise 2 were either 0% or 100%. The ROC curve did not change between Exercises 1 and 2. In Exercise 3, the results of which are not shown here, both CTSs answered with a confidence of either 0% or 100%. This change in interpreting ability should perhaps be described as owing to the influence of training rather than to its effectiveness.

In Exercise 1, the interpreters were instructed to indicate their confidence level regarding the presence of an abnormality on a continuous scale of 0%–100% in order to obtain the data necessary for quantitative determination of diagnostic accuracy by using a ROC curve. From Exercise 2 onward the interpreters were left to their own devices.

The interpretation of thoracic CT scans relates to CT images from a large number of patients. The work of interpretation mainly consists of determining the necessity or otherwise for detailed examination on the basis of an assessment of the presence or absence of abnormalities. Normally, responses do not involve a continuous confidence level. It is therefore possible that the instruction to record a continuous confidence level during the exercises forced a different method on the interpreters from that used in clinical practice.

The confidence distribution obtained in Exercise 1 shown in Figure 5 indicates that the CTSs faithfully followed the instruction to answer on a continuous scale of confidence levels during this exercise, their first experience. From Exercise 2 onward, however, either to alleviate the psychological stress brought on by the unfamiliar act of answering by giving a confidence level, or because they had confidence in their assessment of the presence or absence of an abnormality, it appears that they felt it unnecessary to give a confidence level other than 100% or 0%, as they could either see or not see an abnormality, that one was either present or absent.

This phenomenon may be regarded as the natural result of having learned the basics of diagnosing the presence or absence of an abnormality through the experience of Exercise 1. As shown in the ROC curves in the lower part of Figure 5, however, it should be noted that this sort of change in interpreting ability is not necessarily accompanied by an improvement in diagnostic accuracy. Note: even if the result of the diagnosis of the presence or absence of an abnormality shows an obvious abnormality (100%), the confidence level for qualitative diagnosis of whether this is cancer or another disorder may be around 50%. An answer of 50%, which was not allowed in these experiments, can also not be ruled out.

If this sort of binary action by interpreters is acceptable, the problem is whether a method exists for quantitative evaluation of the results. Currently, the globally popular ROC analysis program[13] cannot handle binary judgment data. Based on the trapezoid method, the ROC program outputs the undervalued Az. It will be necessary to develop a new methodology[14] for drawing ROC curves from the results of binary judgments and determining the Az with high accuracy.

The training carried out immediately prior to Exercise 2 in this report was simple, consisting of CTSs comparing their own result with the right answer. After this training, however, diagnostic accuracy improved significantly. As shown in Figure 6(a), CTSs who achieved high scores before training in Exercise 1 went on to attain high scores after training in Exercise 2, and those with high scores in Exercise 2 also scored highly in Exercise 3 using a CAD system as reference. In contrast, as shown in Figure 6(b), the effect of training (Exercise 2 Az/Exercise 1 Az) was highest for those CTSs with low scores in Exercise 1, and the effect of using the CAD system as reference (Exercise 3 Az/Exercise 2 Az) was also highest for those with the lowest scores following training. The latter trend was also observed in the results of previous exercises carried out with the same CT image database[5]. The scores in Exercise 1, the first experience of CT image interpretation, were almost identical in the previous research and these exercises. The absolute value of the Az scores after using the CAD system as reference, however, was far greater this time than in the previous research. We believe that the reason for this is that the previous experiments included no "training" such as that given in these exercises. This fact suggests that for CTSs to be able to make effective use of the results of a CAD system, they need to learn CT image interpretation methods and acquire a certain level of diagnostic ability.

5. CONCLUSION

To confirm the potential for persons other than physicians to work as CT screeners, we administered CT image interpretation exercises to students of medical technology. As a result, if the students received an appropriate degree of training in image interpretation methods, they achieved a performance on a par with that of working physicians, although this was limited to diagnosing the presence or absence of an abnormality. If they used a CAD system with a degree of complemtarity to the interpreters and appropriate performance their diagnostic accuracy improved even more, demonstrating a further strengthening of their potential as CT screeners.

ACKNOWLEDGEMENTS

We would like to express our gratitude to Drs. Takahiro Uruma, Akira Watanabe, Susumu Kandatsu and students of Tokyo Metropolitan University of Health Sciences. This study was conducted as part of the joint study of Niigata University and Fujitsu Limited. This work was supported in part by a Grant-in-Aid for Cancer Research (15-25) from the Ministry of Health, Labour and Welfare and in part by a Grant-in-Aid for Scientific Research in Priority Areas (15070205) from the Ministry of Education, Science, Sports and Culture in Japan.

REFERENCES

1. T. Matsumoto, T. Miyamoto, T. Suzuki et al. Development of mobile CT unit for lung cancer screening. In *Advances in the Prevention of Occupational Respiratory Diseases*, pp. 485-489. Amsterdam: Elsevier Science, 1998.
2. S. Nawano. Images are surging toward me like a tsunami. *CADM News Letter* 28 (2000): 12-13.
3. N. Nakamura. The exhaustion of radiologists: the necessity of partial withdrawal from image diagnosis. *Radiation Medicine (Japan Radiological Society)* 53: 2003; 131-132. [In Japanese.]
4. A. Matsumoto, ed. *Research and Development of a Spiral CT Lung Cancer Scanning System.* National Institute of Radiological Sciences NIRS-R-50. Tokyo: Jitsugyo-Kohosha, 2003 [In Japanese.]
5. T. Matsumoto, A. Furukawa, K. Machida et al. Methods of evaluating the effectiveness of double-checking in interpreting mass screening images. *Proceedings of SPIE 2004*, vol. 5372: 2004; 1-13.
6. R. M. Friedenberg. The role of the Supertechnologist. *Radiology* 215: 2000; 630-633.
7. M. Matsumoto. *Suupaatekunorojisuto ninteisei kentou iinkai houkokusho* (Report of the Committee to Investigate a Supertechnologist Accreditation System). Tokyo: Japanese Society of Radiological Technology, 2005 (in press). [In Japanese.]
8. K. Hanai. Creating a thoracic CT screener system). Ministry of Health, Labour, and Welfare of Japan Research Cost Subsidies for Cancer Prevention, Etc. Health Sciences Comprehensive Research Project: "Research on Creating a New Testing Model and Improving Testing Efficacy." Tsuchiya Group Materials, December 2004. (Personal communication.) [In Japanese]
9. S. Sone, Z. Li F. Yung et al. Mass screening for lung cancer action project: overall design and findings from baseline screening. *Lancet* 354: 1999; 99-105.
10. T, Matsumoto, A. Furukawa, M. Tsuchikawa et al. Relationship between changes in pupil size over time and diagnostic accuracy. *Proceedings of SPIE 2003*, vol. 5034: 2003; 393-402.
11. S. Yamamoto, H. Takizawa, H. Jiang et al. ACAD system for lung cancer screening test by X-ray CT. In *Proceedings of the 15th International Congress and Exhibition of Computer Assisted Radiology and Surgery*, pp. 605-610. (Excerpta Medica International Congress Series 1230.) Amsterdam: Elsevier Science, 2001.
12. M. Kallergi, G.M. Carne, J. Gavira. Evaluating the performance of detection algorithms in digital mammography. *Med. Phys.* 26 (2): 1999; 267-275.
13. C.E. Metz. *ROCKIT 0.9B Beta Version, IBM compatible ROCKIT User's Guide.* Chicago: University of Chicago Department of Radiology, 1999.
14. T. Matsumoto, N. Fukuda, M. Tsuchikawa et al. Observer performance study for CT-image reading of one slice or multi-slice by the cine display mode of CRT system: An application of the Diagnosis-Dynamic Characteristic (DDC) model, CARS 2001. In *Computer Assisted Radiology and Surgery*, p. 1190. Amsterdam: Elsevier Science, 2001.

資料

胸部CT検診における技師読影（HRCT撮影）の有効性

永野優子　新妻伸二　佐藤和美　帆苅隆　山田一美　稲垣理加

和文要旨

当施設では胸部CT検診の一次検診時に必要があればHRCT撮影を行い、当日のうちに診断をして呼吸器専門機関に紹介をするというシステムをとっている。そのため自発的に技師読影によるHRCT撮影が行われるようになった。そこで技師読影の有効性を検討するために、技師の判断でHRCT撮影を行ったか、または技師がHRCT撮影を行わず読影医師の指示でHRCT撮影を行ったかを調査した。技師読影によるスクリーニング時のHRCT撮影は約9割であり、読影医師の指示でHRCT撮影行ったものは1割程度であった。技師がHRCT撮影を行わずに読影医師の指示でHRCT撮影を行ったものはほとんどが10mm以下の所見であり、11mm以上の所見はほぼ見逃すことなく技師がHRCT撮影を行っていたことがわかった。また、発見肺癌例での前期（平成7年から平成11年）と後期（平成12年から平成16年）との比較では、10mm以下のものでも前期と比べ後期で技師によるHRCT撮影の割合が高くなっており、技師読影の能力の向上がみられた。

キーワード：　胸部CT検診、HRCT、技師読影、肺癌

J Thorac CT Screen 2005;12:307-310

【はじめに】わが国の平均的な胸部CT検診ではスクリーニング時にHigh Resolution CT（以下HRCTと略す）の追加撮影は行われていない。[1,2,3] 多くの検診機関では、一次検診時に検診モードによるCT撮影（時に胸部直接撮影）を行い二重読影・喀痰細胞診の結果、必要があれば後日HRCT撮影を行い診断するというシステムになっている。当施設では一次検診時に必要があるときだけHRCT撮影を行い、当日のうちに診断をして呼吸器専門医療機関に紹介をする、というシステムをとっている。このようなシステム上自然発生的に技師の判断によるHRCT撮影が行われるようになった。そこで技師読影によるHRCT撮影の有効性について検討を試みた。

新潟県労働衛生医学協会
プラーカ健康増進センター
（〒950-0917　新潟県新潟市天神1－1
プラーカ3　5F）

【対象と方法】平成15年8月1日から平成16年8月31日（13ヶ月間）の期間に胸部CT検査を行った2,232人のうち前回ＣＴ検診で有所見者だった例などを除いた、胸部CT検査による肺ドック検診初回受診者・複数回受診者を合わせた1,740人について、技師の判断でHRCT撮影を行ったか、読影医師の指示でHRCT撮影を行ったかを所見別に分類し検討した。

次にCT検診を開始した平成7年からの全期間で発見された肺癌のうちHRCT腫瘍径25mm以下の94例について、平成7年から平成11年までを前期、平成12年から平成16年までを後期とし、技師の判断でCTHR撮影を行ったか、読影医師の指示でHRCT撮影を行ったかをHRCT腫瘍径別に分類し前期と後期とで比較し、技師読影の能力の向上および有効性について検討した。

【結果】調査期間中行われた胸部CT検査による肺ドック検診1,740件中HRCT撮影を行ったのは412件で23.7%にHRCT撮影が行われて

いた。(Table:1)

HRCT撮影を所見別に分類したものでは、技師の判断でHRCT撮影を行ったものについては、Ground Glass Opacity（すりガラス様陰影、以下GGOと略す）：32（7.1%）、小結節：107（23.8%）、肺内リンパ節：61（13.6%）、肺炎または炎症性陰影：35（7.8%）、その他：60（13.4%）、異常なし：103（22.9%）であった。技師がHRCT撮影を行わず第一読影医師の指示でHRCT撮影を行ったものは、GGO:4(8.1%)、小結節:10(20.4%)、肺内リンパ節:9(18.4%)、肺炎または炎症性陰影：3（6.1%）、その他：5（10.2%）、異常なし：18（36.7%）であった。第二読影医師の指摘があり後日HRCT撮影を行ったものは、GGO：1（50%）、小結節：0、肺内リンパ節:1(50%)、肺炎または炎症性陰影：0、その他：0、異常なし：0であった。また、HRCT撮影のうち88.6%に技師の判断でHRCT撮影が行われており、第一読影医師の指示でHRCT撮影を行ったものは10.9%、第二読影医師の指示によるものは0.4%であった。(Fig:1)

全期間中で発見された肺癌については、前期では技師の判断でHRCT撮影を行ったものは10mm以下:7(内GGO7)、11～15mm:10(GGO6)、16～20mm:21（GGO7）、21～25mm:15（GGO4）、技師がHRCT撮影を行わず読影医師の指示でHRCT撮影を行ったものは10mm以下:5(GGO5)、11～15mm：1（GGO1）、16～20mm：2（GGO1）、21～25mm：0、であった。後期では技師の判断でHRCT撮影を行ったものは10mm以下：6（GGO3）、11～15mm：13（GGO）6、16～20mm：5（GGO2）、21～25mm：7（GGO4）、技師がHRCT撮影を行わず読影医師の指示でHRCT撮影を行ったものは10mm以下：2（GGO2）、11～15mm：0、16～20mm：0、21～25mm：0、であった。(Table:2)

【考察】今回調査を行った中で、胸部CT検診のスクリーニング時に技師が有所見と判断しHRCT撮影を行ったものは23.7%と多めであった。そのため平成16年1月23日より被曝線量低減のためHRCT撮影も、体格や部位または所見によって（肺癌が疑われる場合には従来どおり150mAで撮影）低線量の50mA（時に100mA）で撮影を行うこととした。

HRCT撮影の所見別の分類では、技師の判断でHRCT撮影を行ったもので異常なしの件数が103件（26.9%）とかなり多かったが、検診モードだけでは判断できずHRCT撮影を行った結果、血管が異常陰影に見えたもの・葉間肋膜の部分的肥厚がGGO様に見えたものが多く含まれており、実際には診断に必要なHRCT撮影だったと考えられる。技師がHRCT撮影を行わず読影医師からの指示でHRCT撮影を行った結果有所見であったものは33件あったが、そのうち指摘の多かった小結節や肺内リンパ節では10mm以下のものがほとんどであり、11mm以上の所見についてはほぼ見逃すことなくスクリーニング時に技師の判断でHRCT撮影が行われているといえる。また、約9割の所見に対しスクリーニング時にすでに技師の判断でHRCT撮影が行われており読影医師の指示によるHRCT撮影は1割程度であることがわかった。

肺癌例のHRCT撮影については、胸部CT検診を開始した平成7年から平成11年の前期5年間では、技師がHRCT撮影を行わず読影医師の指示でHRCT撮影を行ったものが、HRCT腫瘍径11～15mmに1例、16～20mmに2例の計3例あったが平成12年から平成16年の後期5年間では11mm以上では1例もなく技師が11mm以上の肺癌はほぼ見逃すことなくHRCT撮影を行っていることがわかる。また、10mm以下の肺癌例についても、前期では技師がHRCT撮影を行わず読影医師の指示でHRCT撮影を行ったものが12例中5例（42%）であったが、後期では8例中2例（25%）と技師の判断によるHRCT撮影の割合が高くなっており技師読影の能力の向上が見られた。

以上のことから胸部CT検診のスクリーニング時に技師の判断でHRCT撮影を行うことは有効であるのではないかと考えられる。また読影医師から、所見のある大半の症例がすでにHRCT撮影済みで、安心して診断ができ大変有用であるとの意見をもらっていることを付け加えたい。

【結語】現在シングルスライスCTからマルチスライスCTへの移行期であり、薄いスライスで全肺野を撮影することが可能となったため膨大な数の画像が発生するようになった。今回の調査検討で 10 mm以上の所見についてはほぼ見逃すことなく HRCT 撮影が行われていることがわかり、技師読影によるHRCT撮影は有効であるのではないかと考えられた。この利点をマルチスライス CT 導入時にいかに生かしていけるか今後検討していきたい。

文　献

[1] 西山祥行、三澤潤、田寺守、他：ヘリカルCTを加えた会員制肺がん検診一過去5年間の発見肺がん症例の検討一. 胸部CT検診 1999;6:146-151
[2]川上聡、曽根脩輔、高島庄太夫、他：CT 検診により発見した異型腺腫様過形成の7例. 胸部CT検診 1999;6:259-262
[3] 花岡考臣、曽根脩輔、津島健司、他：2002年度JA長野CT検診車による肺癌1次検診（初年度）のうち、安曇総合病院担当受診者の成績. 胸部CT検診 2003;10:134-138
[4]Sone S, Takashima S, Li F, et al Mass screening for lung cancer with mobile spiral computed tomography scanner. Lancet 351 1998 1242-1245

Efficacy of Radiographer's Reading(Taking HRCT) in Lung Cancer CT screening.

Yuko Nagano, Shinji Niizuma, Kazumi Sato, Takashi Hokari, Kazumi Yamada, Rika Inagaki
Niigata Prefectual Medical Association for Labor Health.

In our clinic, when abnormal findings are observed in screening CT, high resolution computed tomography(HRCT) is additionally performed on the same day, and reflected in the test results. From the beginning of CT mass survey, our radiographers read CT findings on display and took HRCT voluntarily. Now we would like to clarify the efficacy of radiographer's cotribution with taking HRCT and final diagnosis. The results were 1)About 90% of HRCTs were taken by radiographer, only 10% were ordered afterwise by radiologist. 2)The HRCT lesions which were ordered by radiologist showed almost less than 10mm, and all lesions more than 10mm were detected by radiographer. 3)Compared the sizes of disscovered lung cancers in the first 5 years (1995-1999) and the latter 5 years(2000.-2004), detecting rate by radiographer for less than 10mm GGOs and nodular lesions have been improved.

Key words: Lung Cancer CT screening , High-resolution CT, Radiographer's Reading, Lung Cancer
J Thorac CT Screen 2005;12:307-310

Computed tomographic images reflect the biologic behavior of small lung adenocarcinoma: They correlate with cell proliferation, microvascularization, cell adhesion, degradation of extracellular matrix, and K-*ras* mutation

Boming Dong, MD,[a] Masami Sato, MD,[a,b] Akira Sakurada, MD,[a] Motoyasu Sagawa, MD,[a,c] Chiaki Endo, MD,[a] Shulin Wu, MD,[a] Sumitaka Yamanaka, MD,[a] Akira Horii, MD,[d] and Takashi Kondo, MD[a]

GTS

Background: We previously reported that the computed tomographic M/L ratio (area of the tumor in the mediastinal computed tomographic image/area of the tumor in the lung computed tomographic image) of small peripheral lung adenocarcinoma is correlated with patient prognosis.

Methods: Immunostaining for p53, bcl-2, Ki-67, vascular endothelial growth factor, CD34, matrix metalloproteinase 2, matrix metalloproteinase 9, tissue inhibitor of matrix metalloproteinase 2, and mutation of K-*ras* was assessed in 131 surgically resected, primary peripheral lung adenocarcinomas of 30 mm or less in maximum diameter to clarify the relationship between computed tomographic findings and biologic activities.

Results: The numbers of patients with high labeling indexes of Ki-67 and high expression of vascular endothelial growth factor, CD34, matrix metalloproteinase 2, and matrix metalloproteinase 9 in the solid-type group (computed tomographic M/L ratio ≥50%) were significantly higher than those in the faint density–type group (computed tomographic M/L ratio <50%; $P = .04$ for Ki-67, $P = .03$ for vascular endothelial growth factor, $P = .0009$ for CD34, $P = .001$ for matrix metalloproteinase 2, and $P = .00001$ for matrix metalloproteinase 9). The number of patients with high levels of CD44v6 or tissue inhibitor of matrix metalloproteinase 2 staining in the faint density–type group was significantly higher than that in the solid-type group ($P = .02$ for CD44v6 and $P = .01$ for tissue inhibitor of matrix metalloproteinase 2). Independent variables capable of predicting computed tomographic M/L ratio included CD34, matrix metalloproteinase 2, matrix metalloproteinase 9, and tissue inhibitor of matrix metalloproteinase 2 ($P = .0093$, $P = .0003$, $P = .0027$, and $P = .01$, respectively; binary logistic regression analysis).

Conclusions: Our results suggest that the computed tomographic image of small lung adenocarcinoma is correlated with biologic activities and thus provides possible prognostic information.

From the Department of Thoracic Surgery,[a] Institute of Development, Aging and Cancer, Tohoku University School of Medicine, Sendai, Japan; the Department of Thoracic Surgery,[b] Miyagi Cancer Center, Natori, Japan; the Department of Thoracic Surgery,[c] Kanazawa Medical University, Kanazawa, Japan; and the Department of Molecular Pathology,[d] Tohoku University, Sendai, Japan.

This work was supported in part by the Grant-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

Received for publication March 1, 2005; revisions received April 29, 2005; accepted for publication May 16, 2005.

Address for reprints: Masami Sato, MD, Department of Thoracic Surgery, Miyagi Cancer Center, 47-1, Medeshima-Shiote-aza-Nodayama, Natori, 981-1293, Miyagi, Japan (E-mail: m-sato@mcc.pref.miyagi.jp).

J Thorac Cardiovasc Surg 2005;130:733-9

0022-5223/$30.00

doi:10.1016/j.jtcvs.2005.05.013

Recently, the number of patients with small peripheral lung cancer detected by means of computed tomographic (CT) scanning has been increasing. Despite the small size of these tumors, some of these patients die of cancer recurrence. Several factors, including lymph node metastasis, histologic subtype, size of the central fibrosis, and bronchioloalveolar carcinoma (BAC) component, have been proved to be of prognostic importance, but few such factors can be evaluated preoperatively.[1-4] In our previous report we reviewed the CT images of 143 patients with primary peripheral lung adenocarcinoma of 30 mm or less in maximum diameter and classified the patients according to the CT M/L ratio (area

of tumor in the mediastinal CT image/area of tumor in the lung CT image) into a faint density–type group (CT M/L ratio <50%) and a solid-type group (CT M/L ratio ≥50%).[5] As a result, we found that the 5-year survival of patients with faint density–type CT images was significantly better than that of patients with solid-type CT images. Even when the analyses were limited to the patients without nodal involvement, there was also a significant difference between the 2 groups. Multivariate analysis revealed the effect on prognostic influence of the CT M/L ratio on survival to be the second highest after that of the N factor.[5] Thus preoperative classification of patients according to the CT M/L ratio is a simple but powerful predictor of prognosis for patients with small-sized lung adenocarcinoma. Adjuvant treatments, such as chemotherapy and molecular target therapy, are probably necessary for patients with sold-type CT images. Thus it is important to know the mechanism by which CT M/L ratio can be used to predict the prognosis for patients with small-sized lung adenocarcinoma.

The CT M/L ratio might be affected by the pathologic structure of the tumor. BAC components show low cellular density because the tumor cells often extend by mimicking normal alveolar structure. In contrast, fibroblastic proliferation, formation of a central scar, and neovascularization might all contribute to an increase in the value of the CT M/L ratio. None of these factors, however, seems to uniquely predominate in this CT-pathology-survival correlation. On the other hand, p53, Ki-67, bcl-2, CD44v6, vascular endothelial growth factor (VEGF), matrix metalloproteinase 2 (MMP-2), MMP-9, tissue inhibitor of MMP-2 (TIMP-2), and K-*ras*, which were reported to be prognostic factors of small adenocarcinoma, have some influence on histologic structure by affecting cell proliferation, apoptosis, microvascularization, cell adhesion, and the degradation of the extracellular matrix.[6-10] The purpose of this study was to identify the biologic factors that either correlated with the CT M/L ratio or predict the prognosis of small lung adenocarcinoma, so as to clarify the mechanism of the CT-pathology-survival correlation.

Materials and Methods

Patients, CT Image Analyses, and Tissue Specimen

In this study, we analyzed resected tumor specimens from 131 patients with primary peripheral lung adenocarcinomas with diameters of 30 mm or less in a period between January 1990 and December 1993. Of the 131 patients, 126 underwent lobectomies, 3 underwent pneumonectomies, and 2 underwent segmental resections combined with systematic hilar and mediastinal node dissection. The CT M/L ratios of tumors were calculated as previously described,[5] and all the patients were classified into one of two groups: a solid-type tumor group and a faint density–type tumor group. The medical records of all the patients were reviewed for clinical and pathologic characteristics. The length of survival was defined as the period from the day of the operation to the last day of follow-up or the date of death from any cause. Thirty 3-μm–thick sections and two 5-μm–thick sections were cut from paraffin blocks of the resected tumor for immunohistochemical staining and microdissection, respectively.

Immunohistochemistry

Immunohistologic staining was performed by the streptoavidin-biotin method (Histofine SAB-PO Kit; Nichirei, Tokyo, Japan), followed by counterstaining with hematoxylin. For antigen retrieval, sections were autoclaved in 10 mmol/L citrate buffer (pH 6.0) before staining. Concentrations of the primary antibodies were optimized by preliminary examinations. Staining without the primary antibody was routinely performed as a negative control. Known immunostaining positive lung adenocarcinoma or squamous cell carcinoma tissue specimens were used as positive controls. The antibodies used in this study were monoclonal p53 antibody (clone DO-7, DAKO, Glostrup, Denmark; diluted 1:50), monoclonal Bcl-2 antibody (clone 124, DAKO, diluted 1:50), monoclonal Ki-67 antibody (clone MIB-1, DAKO, diluted 1:50), monoclonal CD44v6 antibody (clone 2F-10, R&D systems, Minneapolis, Minn; diluted 1:250), monoclonal VEGF antibody (clone R11, IBL, Fujioka, Gunma, Japan; diluted 1:50), monoclonal CD34 antibody (clone QBEnd 10, DAKO, diluted 1:80), monoclonal MMP-2 antibody (clone 8B4, Santa Cruz Biotechnology, Santa Cruz, Calif; diluted 1:200), monoclonal MMP-9 antibody (clone 2C3, Santa Cruz Biotechnology, diluted 1:250), and monoclonal TIMP-2 antibody (clone 3A4, Santa Cruz Biotechnology, diluted 1:250). All the slides were reviewed by three physicians without any knowledge of clinical outcomes or any other clinicopathologic data: two of them (M.S. and A.S.) received formal cytopathologic training. Microscopic evaluation was performed by counting more than 1000 tumor cells in more than 5 randomly selected high-power fields (400×) from different representative parts of the tumor. The Ki-67 labeling index (percentage of nuclear-stained cells) was defined as high when it exceeded the mean value (Figure 1, *a*).[6] The p53[11] staining was defined as positive when more than 10% of the tumor cells showed nuclear staining (Figure 1, *b*). Bcl-2 (Figure 1, *c*),[6] CD44v6 (Figure 1, *d*),[7] MMP-2 (Figure 1, *e*),[12] MMP-9 (Figure 1, *f*),[12] and TIMP-2 (Figure 1, *g*)[12] were defined as positive when more than 10% of the tumor cells showed cytoplasmic staining. For evaluation of VEGF, immunostaining was used to determine the percentage of immunoreactive cells, with the cutoff point for distinguishing specimens with low from high VEGF expression set at 30% of carcinoma cells (Figure 1, *h*). Microvessel density (MVD), as measured by CD34 immunostaining, was determined by using a modification of the technique described by Weidner and colleagues (Figure 1, *i*).[13] The mean MVD was used as a cutoff point for distinguishing between tumors with low and high MVDs.

Mutation Analysis of K-*ras* at Codons 12 and 13

Two hundred or more tumor cells were laser captured on separate Caps by using a laser capture microdissection microscope (Leica AS LMD; Leica Microsystems, Tokyo, Japan), according to the manufacturer's instructions. DNA was extracted with the QIAamp DNA Mini Kit (50; QIAGEN, Tokyo, Japan), according to the protocol provided by the manufacturer. Exon 1 of the K-*ras* gene was amplified by the nested polymerase chain reaction (PCR) with the first

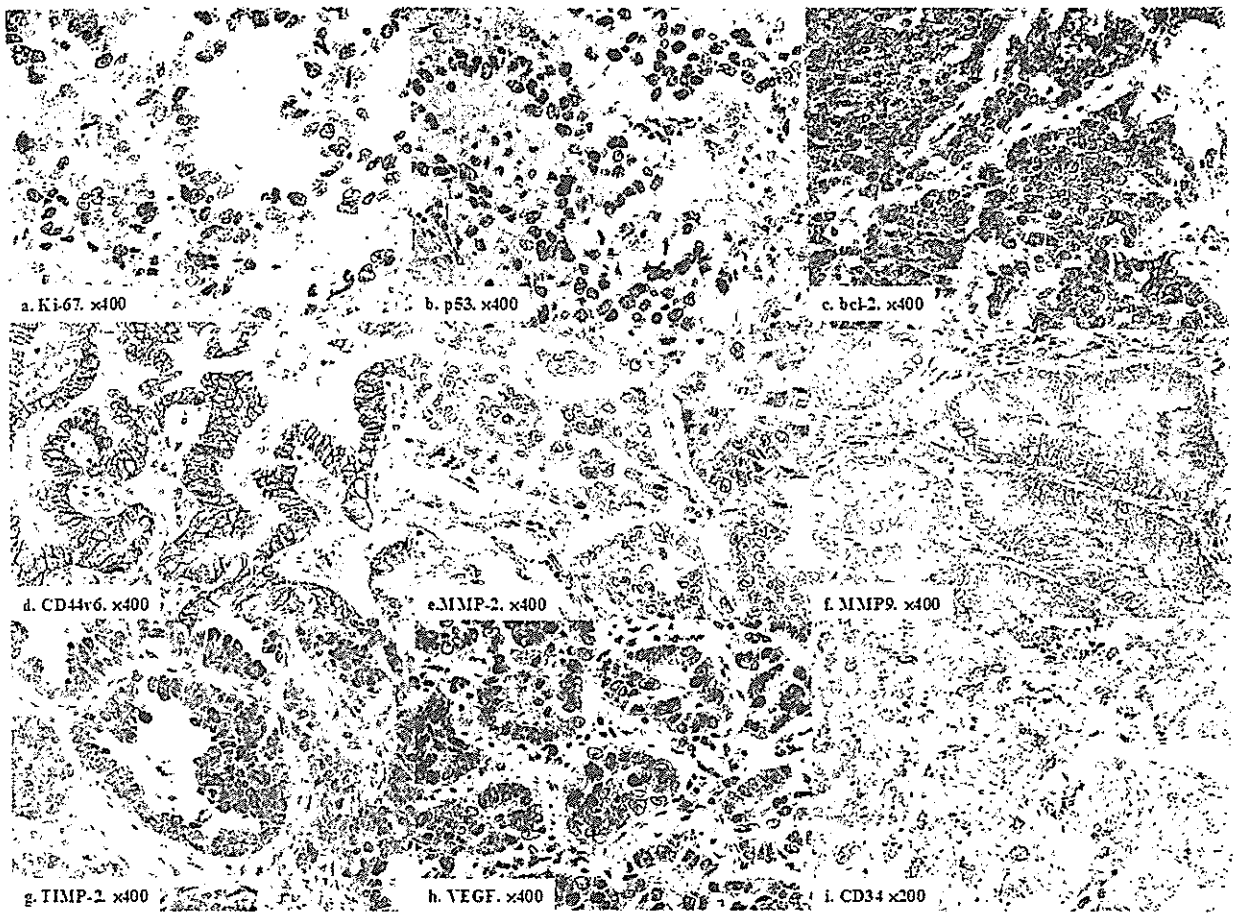

Figure 1. Immunohistochemical staining of serial sections with various antibodies. Nuclear accumulation of Ki-67 (a) and p53 (b) proteins and cytoplasmic accumulation of bcl-2 (c), CD44v6 (d), MMP-2 (e), MMP-9 (f), TIMP-2 (i), and VEGF (h) proteins are shown in neoplastic cells. (Original magnification 400×.) The CD34-stained endothelial cell clusters were considered as a single countable microvessel, and a tumor area with a high microvessel count is shown in i. (Original magnification 200×.)

primer set (5′-GACATGTTCTAATATAGTCACAT-3′ and 5′-GTCCTGCACCAGTAATATGC-3′) and the nested primer set (5′-AGGCCTGCTGAAAATGACTG-3′ and 5′-CCTCTATTGTTTG-GATCATATTC-3′), which yielded 205-bp and 125-bp DNA fragments, respectively, covering codons 12 and 13. PCR was carried out in a final 25-μL reaction mixture containing 4.5 mmol/L Tris-HCl (pH 8.8), 67 mmol/L $(NH_4)_2SO_4$, 6.7 mmol/L β-mercaptoethanol, 4.5 μmol/L ethylenediamine tetraacetic acid, 4.5 mmol/L $MgCl_2$, 0.75 pmol each of deoxyribonucleotide triphosphates, 2.5 pmol of each primer, and 0.25 units of Taq DNA polymerase.[14] The mixture was heated at 95°C for 5 minutes and then subjected to 40 PCR cycles (94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 30 seconds). The resulting nested PCR products were run on 4% agarose gels, the 125-bp fragments were isolated, and the DNAs were purified for sequencing reactions. Nucleotides of both the sense and antisense strands were determined by using the BigDye Terminator Cycle Sequencing Ready FS Reaction Kit (Applied Biosystems, Foster City, Calif) and ABI PRISM 310 Genetic Analyzer (Applied Biosystems) by means of methods described previously.[15] This study was approved by the Ethical Committee of Tohoku University School of Medicine and conducted according to the Declaration of Helsinki principles. The appropriate procedure for obtaining informed consent was followed for all the probable individuals participating in the study. Permission for investigations on the old specimens with informed consent that was unavailable during the period of study was obtained from the Ethical Committee of Tohoku University School of Medicine on the basis of a strict evaluation about the essentiality and necessity of the research.

Statistical Analysis

Survival was calculated by means of the Kaplan-Meier method, and statistical analysis was performed by the log-rank test. The χ^2

test was used to examine group demographic differences. Binary logistic regression analysis was used to identify predictors of CT M/L ratio. The Cox proportional hazards model with a forward stepwise procedure was applied for multivariate analysis. All statistical analyses were performed with version 5.0 of the StatView software package (SAS Institute, Inc, Cary, NC).

Results

Patient characteristics are summarized in Table 1. There were 53 patients with solid-type tumors and 78 patients with faint density–type tumors. There were no significant differences in sex, age distribution, nodal involvement, or histologic grade between the 2 groups. The 5-year survival of patients with faint density–type images was 81.6%, which was significantly higher than that of patients with solid-type CT images (5-year survival, 63.8%; $P = .004$; Figure 2, *a*).

Positive staining for Ki-67 and p53 protein was detected in the nuclei (Figure 1, *a* and *b*). In all cases the percentage of tumor cells positive for Ki-67 protein ranged from 0.6% to 78.4%, and the mean Ki-67 labeling index was 18.2% ± 16.2%. The expressions of bcl-2, CD44v6, MMP-2, MMP-9, TIMP-2, and VEGF proteins were generally detected in the cytoplasm of tumor cells (Figure 1, *c-h*). The mean MVD stained by CD34 (Figure 1, *i*) for all patients was 59.2 ± 33.8. The numbers of patients with high labeling indexes of Ki-67 and high expressions of VEGF, CD34, MMP-2, and MMP-9 in the solid-type group were significantly greater than those in the faint density–type group ($P = .04$ for Ki-67, $P = .03$ for VEGF, $P = .0009$ for CD34, $P = .001$ for MMP-2, and $P = .00001$ for MMP-9; Table 1). The numbers of patients with high levels of CD44v6 or TIMP-2 staining in the faint density–type group were significantly greater than those in the solid-type group ($P = .02$ for CD44v6 and $P = .01$ for TIMP-2, Table 1). There were no significant differences in the expressions of p53 and bcl-2 proteins between the 2 groups ($P = .64$ for p53 and $P = .10$ for bcl-2, Table 1). The 5-year survivals of patients with high indexes of Ki-67 or low levels of TIMP-2 staining were poorer than those of patients with low indexes of Ki-67 or high levels of TIMP-2 staining ($P = .026$ for Ki-67 and $P = .0005$ for TIMP-2; Figure 2, *b* and *c*). The expressions of VEGF, CD34 (MVD), MMP-2, and MMP-9 proteins seem to be unfavorable prognostic factors (Table 1).

DNA extraction, amplification of K-*ras* by nested PCR, and direct sequencing was successful in 95 cases (43 solid type and 52 faint density type), and K-*ras* mutations were found in 18 (19%) tumors, including 17 mutations on codon 12 (9 GGTGly to TGTCys, 4 GGTGly to AGTSer, and 4 GGTGly to GATAsp) and 1 mutation on codon 13 (GGCGly to GACAsp). The frequency of K-*ras* mutations in the solid-type group was significantly higher than that in the faint density–type group (Table 1). The 5-year survival of patients without a K-*ras* mutation was 0.78, which was significantly better than that of those with a K-*ras* mutation, which was 0.62 ($P = .04$; Figure 2, *d*).

Binary logistic regression analysis revealed the independent variables of significance in predicting CT M/L ratio to be CD34, MMP-9, MMP-2, and TIMP-2 ($P = .0093$, $P = .0003$, $P = .0027$, and $P = .01$, respectively). The multivariate analysis showed N factor, TIMP-2, Ki-67, and tumor size to be significant prognostic factors ($P < .0001$, $P = .0015$, $P = .030$, and $P = .048$, respectively; Table 2).

Discussion

In this study we found that patients with solid-type tumors had significantly poorer prognoses than those with faint density–type tumors (Figure 2). We also observed that a poor prognosis was associated with high Ki-67 expression and K-*ras* mutation (Figure 2). In multivariate analysis, Ki-67 was proved to be an independent prognostic factor (Table 2). These results were in good agreement with previous reports.[6,11,16-20] In the solid-type group, the number of patients with high labeling indexes of Ki-67, as well as K-*ras* mutation, was significantly higher than that in the faint density–type group (Table 1). Positive correlations between K-*ras* mutation and a high Ki-67 labeling index, a high MVD, a tumor diameter of greater than 20 mm, or low expression of CD44v6 (results not shown) suggest that the K-*ras* mutation affects biologic behaviors, such as cell proliferation, microvascularization, and cell adhesion, and thus the CT M/L ratio in small peripheral lung adenocarcinoma.

Recently, some researchers have reported that increased levels of VEGF expression and new vessel formation were associated with the poorer survival of patients with non–small cell lung cancer.[8,9,21-23] In this study, the numbers of patients with high expression of VEGF and high MVD in the solid-type group were significantly greater than those in the faint density–type group (Table 1). In binary logistic regression analysis, CD34 (MVD) was an independent variable of significance in predicting CT M/L ratio. All these facts indicated that angiogenesis in the solid-type group was much more vigorous than that in the faint density–type group. The newly formed vessels contribute to the increasing CT M/L ratio, as well as to the patient's poor prognosis. Molecular target therapy against tumor-related vessel formation might be adequate for patients with solid-type CT images.

Several recent reports have confirmed that MMP-2 and MMP-9 might predict the outcome of non–small cell lung cancer.[9,24-26] Kumaki and colleagues[27] and Nawrocki and associates[28] have also reported that the expressions of MMP-2, MMP-9, and TIMP-2 differed between atypical adenomatous hyperplasia and BAC or between BAC and invasive adenocarcinoma at the mRNA level or as shown by immunostaining study. Therefore, we investigated the expression of these 3 factors and found that the numbers of patients with high levels of expression of MMP-2 and

TABLE 1. Clinicopathologic characteristics and summary of immunohistochemical studies and K-*ras* mutations

	CT findings					
	No. of patients	Solid type*	Faint density type†	χ^2 *P* value	5-y survival	Log-rank test *P* value
Sex				.91		.40
Male	61	25	36		76.4	
Female	70	28	42		72.5	
Age (y)	62.9 ± 9.7	62.9 ± 9.6	63.0 ± 9.8	.54		.94
Tumor size				.03		.01
≤20 mm	52	15	37		83.9	
>20 mm	79	38	41		68.8	
Nodal involvement				.09		<.00001
N0	99	36	63	N^+ vs N^-		N^+ vs N^-
N1	7	2	5			
N2	25	15	10			
Histologic grade				.015		.02
Well	43	13	30	Well vs moderate + poor		Well vs moderate + poor
Moderate	63	29	34			
Poor	19	5	14			
p stage				.08		<.00001
I	96	35	61	Stage I vs stage II-IV		Stage I vs stage II-IV
II	8	2	6			
III	26	16	10			
IV	1	1	0			
p53				.64		.16
Positive	66	28	38		67.9	
Negative	65	25	40		81.3	
Ki-67				.04		.03
High	48	25	23		67.7	
Low	83	28	55		79.2	
Bcl-2				.10		.051
Positive	40	12	28		82.3	
Negative	91	41	50		70.9	
CD44v6				.02		.056
Positive	80	26	54		79.4	
Negative	51	27	24		66.4	
VEGF				.03		.18
Positive	50	26	24		66.7	
Negative	81	27	54		80.4	
MVD (CD34)				.0009		.23
High	54	31	23		69.9	
Low	77	22	55		77.8	
MMP-2				.01		.82
Positive	62	34	28		75.0	
Negative	69	19	50		74.2	
MMP-9				.0001		.058
Positive	93	47	46		69.0	
Negative	38	6	32		87.5	
TIMP-2				.01		.0005
Positive	90	30	60		83.3	
Negative	41	23	18		55.8	
K-*ras* mutations				.0003		.04
Mutant	18	15	3		62.2	
Wild type	77	28	49		78.4	

CT, Computed tomography; *VEGF*, vascular endothelial growth factor; *MVD*, microvessel density; *MMP-2*, matrix metalloproteinase 2; *MMP-9*, matrix metalloproteinase 9; *TIMP-2*, tissue inhibitor of matrix metalloproteinase 2. *Solid type: area of the tumor in mediastinal window/area of the tumor in lung window of 50% or greater. †Faint density type: area of the tumor in mediastinal window/area of the tumor in lung window of less than 50%.

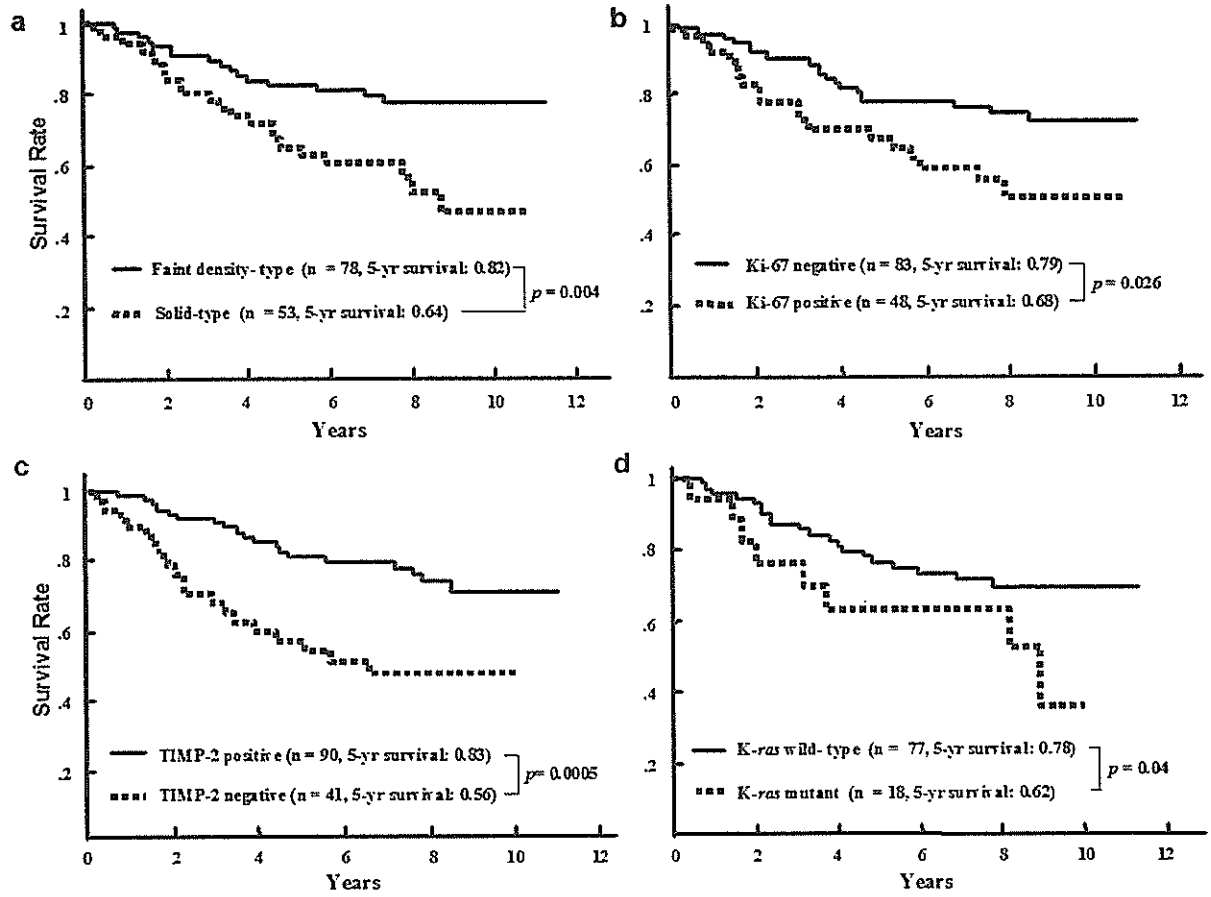

Figure 2. Survivals of patients with faint density–type images, negative Ki-67 immunostaining, or positive TIMP-2 immunostaining were significantly greater than those of patients with solid-type images, positive Ki-67 immunostaining, or negative TIMP-2 immunostaining ($P = .004$, $P = .026$, and $P = .0005$ in a, b, and c, respectively, by using the log-rank test). The 5-year survival of patients without K-*ras* mutation was 0.78 compared with 0.62 for patients with K-*ras* mutations ($P = .04$ in d).

MMP-9 or low levels of expression of TIMP-2 in the solid-type group were significantly greater than those in the faint density–type group (Table 1). Furthermore, MMP-2, MMP-9, and TIMP-2 were proved to be the independent variables of significance in predicting CT M/L ratio, and TIMP-2 was proved to be a strong independent prognostic factor of small lung adenocarcinoma both in univariate and multivariate analysis (Tables 1 and 2 and Figure 2, *c*). All these facts suggested that in addition to angiogenesis, degradation of the extracellular matrix was another important factor that affects the CT M/L ratio, and TIMP-2 might have an important role in the development and extension of lung adenocarcinoma. Further research is necessary to explore the mechanisms of TIMP-2, which affect the CT M/L ratio and the prognosis of small lung adenocarcinoma.

TABLE 2. Multivariate analyses of prognostic factors

Variables	Estimated coefficient	SE	*P* value	Relative risk (95% CI)
Nodal involvement (N0 vs N1 + 2)	−2.53	0.36	<.0001	0.08 (0.04-0.16)
TIMP-2	1.06	0.33	.0015	2.88 (1.50-5.54)
Ki-67	−0.72	0.33	.030	0.49 (0.25-0.93)
Tumor size (>20 mm vs ≤20 mm)	−0.78	0.39	.048	0.46 (0.21-0.99)

SE, Standard error; *CI*, confidence interval; *TIMP-2*, tissue inhibitor of matrix metalloproteinase 2.

In conclusion, tumors with solid-type or faint density–type CT images show significant correlations with factors such as immunostaining of Ki-67, CD44v6, VEGF, CD34 (MVD), MMP-2, MMP-9, TIMP-2, and K-*ras* mutation;

several factors probably affect the promotion of malignant potential in solid-type tumors. To the best of our knowledge, our study is the first to evaluate the correlation of conventional CT images with these biologic factors and to indicate that the CT M/L ratio might provide some information on a tumor's biologic behaviors and might be a prognostic tool for evaluating patients with small peripheral lung adenocarcinoma in the pretreatment period.

We thank Dr Barbara Lee Smith Pierce (University of Maryland University College) for editorial work in the preparation of this article.

References

1. Suzuki K, Nagai K, Yoshida J, Nishimura M, Takahashi K, Yokose T, et al. Conventional clinicopathologic prognostic factors in surgically resected non-small cell lung carcinoma. *Cancer.* 1999;86:1976-84.
2. Noguchi M, Morikawa A, Kawasaki M, Matsuno Y, Yamada T, Hirohashi S, et al. Small adenocarcinoma of the lung—histologic characteristics and prognosis. *Cancer.* 1995;75:2844-52.
3. Suzuki K, Yokose T, Yoshida J, Nishimura M, Takahashi K, Nagai K, et al. Prognostic significance of the size of central fibrosis in peripheral adenocarcinoma of the lung. *Ann Thorac Surg.* 2000;69:893-7.
4. Higashiyama M, Kodama K, Yokouchi H, Takami K, Mano M, Kido S, et al. Prognostic value of bronchiolo-alveolar carcinoma component of small lung adenocarcinoma. *Ann Thorac Surg.* 1999;68:2069-73.
5. Dong B, Sato M, Sagawa M, Endo C, Usuda K, Sakurada A, et al. Computed tomographic image comparison between mediastinal and lung windows provides possible prognostic information in patients with small peripheral lung adenocarcinoma. *J Thorac Cardiovasc Surg.* 2002;124:1014-20.
6. Ishida H, Irie K, Itoh T, Furukawa T, Tokunaga O. The prognostic significance of p53 and bcl-2 expression in lung adenocarcinoma and its correlation with Ki-67 growth fraction. *Cancer.* 1997;80:1034-45.
7. Ramasami S, Kerr MK, Chapman DA, King G, Cockburn JS, Jeffrey RR. Expression of CD44v6 but not E-cadherin or β-catenin influences prognosis in pulmonary adenocarcinoma. *J Pathol.* 2000;192:427-32.
8. Shibusa T, Shijubo N, Abe S. Tumor angiogenesis and vascular endothelial growth factor expression in stage I lung adenocarcinoma. *Clin Cancer Res.* 1998;4:1483-7.
9. Kodate M, Kasai T, Hashimoto H, Yasumoto K, Iwata Y, Manabe H. Expression of matrix metalloproteinase (gelatinase) in T1 adenocarcinoma of the lung. *Int J Pathol.* 1997;47:461-9.
10. Silini ME, Bosi F, Pellegata Gino Volpato SN, Romano A, Tinelli SNC, Ranzani Enrico Solcia NG, et al. K-*ras* gene mutations: an unfavorable prognostic marker in stage I lung adenocarcinoma. *Virchows Arch.* 1994;424:367-73.
11. Minami K, Saito Y, Imamura H, Okamura A. Prognostic significance of p53, ki-67, VEGF and Glut-1 in resected stage I adenocarcinoma of the lung. *Lung Cancer.* 2002;38:51-7.
12. Michael M, Babic B, Khokha R, Tsao M, Ho J, Pintilie M, et al. Expression and prognostic significance of metalloproteinases and their inhibitors in patients with small-cell lung cancer. *J Clin Oncol.* 1999;17:1802-8.
13. Weidner N, Semple JP, Welch WR, Folkman J. Tumor angiogenesis and metastasis—correlation in invasive breast carcinoma. *N Engl J Med.* 1991;324:1-8.
14. Kimura M, Abe T, Sunamura M, Matsuno S, Horii A. Detailed deletion mapping on chromosome arm 12 in human pancreatic adenocarcinoma: identification of a 1-cM region of common allelic loss. *Genes Chromosomes Cancer.* 1996;17:88-93.
15. Sakurada A, Suzuki A, Sato M, Yamakawa H, Orikasa K, Uyeno S, et al. Infrequent genetic alterations of the PTEN/MMAC1 gene in Japanese patients with primary cancers of the breast, lung, pancreas, kidney and ovary. *Jpn J Cancer Res.* 1997;88:1025-8.
16. Takahashi S, Kamata Y, Tamo W, Koyanagi M, Hatanaka R, Yamada Y, et al. Relationship between postoperative recurrence and expression of cyclin E, p27, and Ki-67 in non-small cell lung cancer without lymph node metastases. *Int J Clin Oncol.* 2002;7:349-55.
17. Demarchi LMMF, Reis MM, Palomino SAP, Takagaki CFTY, Beyruti R, Saldiva PHN, et al. Prognostic values of stromal proportion and PCNA Ki-67, and p53 proteins in patients with resected adenocarcinoma of the lung. *Mod Pathol.* 2000;13:511-20.
18. Huncharek M, Muscat J, Geschwind JF. K-*ras* oncogene mutation as a prognostic marker in non-small cell lung cancer: combined analysis of 881 cases. *Carcinogenesis.* 1999;20:1507-10.
19. Fukuyama Y, Mitsudomi T, Sugio K, Ishida T, Akazawa K, Sugimachi K. K-*ras* and p53 mutations are an independent unfavourable prognostic indicator in patients with non-small-cell lung cancer. *Br J Cancer.* 1997;75:1125-30.
20. Cho JY, Kim JH, Lee YH, Chung KY, Kim SK, Gong SJ, et al. Correlation between K-*ras* gene mutation and prognosis of patients with non-small cell lung carcinoma. *Cancer.* 1997;79:462-7.
21. Folkman J. Angiogenesis in cancer, vascular, rheumatoid and other disease. *Nat Med.* 1995;1:27-31.
22. Yano T, Tanikawa S, Fujie T, Masutani M, Horie T. Vascular endothelial growth factor expression and neovascularisation in non-small cell lung cancer. *Eur J Cancer.* 2000;36:601-9.
23. Inoshima N, Nakanishi Y, Minami T, Izumi M, Takayama K, Yoshino I, et al. The influence of dendritic cell infiltration and vascular endothelial growth factor expression on the prognosis of non-small cell lung cancer. *Clin Cancer Res.* 2002;8:3480-6.
24. Pinto AC, Carvalho EOP, Antonangelo L, Garippo A, Da Silva AG, Soares F, et al. Morphometric evaluation of tumor matrix metalloproteinase 9 predicts survival after surgical resection of adenocarcinoma of the lung. *Clin Cancer Res.* 2003;9:3098-104.
25. Thomas P, Khokha R, Shepherd AF, Feld R, Tsao MS. Differential expression of matrix metalloproteinases and their inhibitors in non-small cell lung cancer. *J Pathol.* 2000;190:150-6.
26. Cox G, Joes JL, Andi A, Waller DA, O'Byrne KJ. A biological staging model for operable non-mall cell lung cancer. *Thorax.* 2001;56:561-6.
27. Kumaki F, Matsui K, Kawai T, Ozeki Y, Yu ZX, Ferrans JV, et al. Expression of matrix metalloproteinases in invasive pulmonary adenocarcinoma with bronchioloalveolar component and atypical adenomatous hyperplasia. *Am J Pathol.* 2001;159:2125-34.
28. Nawrocki B, Polette M, Marchand V, Monteau M, Gillery P, Tournier JM, et al. Expression of matrix metalloproteinases and their inhibitors in human bronchopulmonary carcinomas: quantitative and morphological analyses. *Int J Cancer.* 1997;72:556-64.

低線量CT肺癌検診の有効性評価

中山富雄[1]・鈴木隆一郎[1]

[1]大阪府立成人病センター調査部

Japanese Journal of Lung Cancer

肺癌 第46巻 第7号 2006年12月

低線量CT肺癌検診の有効性評価

中山富雄[1]・鈴木隆一郎[1]

要旨 —— **目的.** 低線量CTは，肺癌発見率の向上・発見肺癌の腫瘍径の小型化・I期例の増加・高い生存率という点で注目されている．しかしこのような症例単位での研究は，偏りの影響を受けやすく，癌検診の評価としては，癌死亡率をエンドポイントとした疫学研究が必要である．**方法と結果.** Japan Lung Cancer Screening Study (JLCSS) は，肺癌死亡率をエンドポイントとするコホート研究である．1995年から行われたCT検診の受診者46,733人と単純X線検診受診者91,970人を登録し，2002年まで追跡している．**結論.** この研究結果に加えて，高い要精検率・高い費用・放射線被曝という三つの不利益を解消することが，低線量胸部CT検診にとっての課題であり，解決されない状況での普及は推奨できない．(肺癌. 2006;46:871-876)

索引用語 —— 肺癌検診，低線量CT，死亡率

The Evaluation of the Effectiveness of Low Dose Helical Computed Tomography Screening

Tomio Nakayama[1]*; Takaichiro Suzuki*[1]

ABSTRACT —— ***Objective.*** The low dose helical computed tomography (LDCT) screening for lung cancer is attracting attention because of its high detection rate, ability to detect small tumors, and high survival rate of detected cases. However, an epidemiologic study in which the endpoint is cancer mortality is necessary to evaluate the effectiveness of cancer screening because these case studies have easily contaminated biases. ***Method and Result.*** Japan Lung Cancer Screening Study (JLCSS) is a cohort study that compared the lung cancer mortality rates of CT screened group and chest X-ray screened group as an endpoint. Since 1995, 46,733 people registered in the CT screening group and 91,970 people registered in the chest X-ray screening group, and they were followed up until 2002. ***Conclusion.*** There is a problem to cancel three disadvantages such as high dose examination rate, cost and radiation exposure in addition to the results of this study for LDCT screening. The spread of the LDCT screening for lung cancer cannot be recommended until this problem is solved. (*JJLC.* 2006;46:871-876)

KEY WORDS —— Lung cancer screening, Helical CT, Mortality

1. はじめに

低線量CT検診は，1993年に「東京から肺癌をなくす会」で開始されて以来，[1] 肺癌対策の切り札として，我が国ばかりではなく世界でも大変注目されている．国内では研究ばかりではなくすでに，府県あるいは市町村の事業としてCT検診を実施する地区も見られてきた．肺癌検診の今後の展開を検討する上で，CT検診のevidenceを整理し，現状の方向性について検討するものである．

[1]大阪府立成人病センター調査部．
別刷請求先：中山富雄．大阪府立成人病センター調査部疫学課．〒537-8511 大阪市東成区中道1-3-3(e-mail: nakayama-to@mc.pref.osaka.jp)．
[1]Department of Cancer Control and Statistics, Osaka Medical Center for Cancer and Cardiovascular Diseases, Japan.
Reprints: Tomio Nakayama, Department of Cancer Control and Statistics, Osaka Medical Center for Cancer and Cardiovascular Diseases, 1-3-3 Nakamichi, Higashinari-ku, Osaka 537-8511, Japan (e-mail: nakayama-to@mc.pref.osaka.jp).

Table 1. Summary of the Low-Dose Helical CT Screening

		ALCA	Nagano	ELCAP	Hitachi
Study participant age (years old) smoking		40-79 smoker	40-	60- smoker	50-69
Beginning of the study		1993	1996	1993	1998
First screening	Participants	1,611	5,483	1,000	7,956
	Cancer detected	14	23	27	36
	Detection rate (%)	0.87	0.42	2.7	0.44
	Mean tumor size (mm)	19.8	15.1	13.8	17.0
	Stage IA (%)	71	91	81	78
	5-year survival (%)	76.2	-	-	-
Repeated screening	Total participants	7,891	8,303	1,184	5,568
	Cancer detected	22	37	7	4
	Detection rate (%)	0.28	0.45	0.59	0.07
	Mean tumor size (mm)	14.6	12.0	12.1	16.0
	Stage IA (%)	82	86	71	100
	5-year survival (%)	64.9	-	-	-

ALCA: Anti-Lung Cancer Association, ELCAP: Early Lung Cancer Action Project.

2. CT 検診に関する過去の報告

Table 1 に代表的な CT 検診に関する報告をまとめた.[1-4] 対象となる集団の性・年齢・喫煙状況・人種等が異なるため発見率等は異なるが，今までの報告に共通したことは，①初回の CT 検診の癌発見率は胸部単純 X 線検診の数倍に相当する，②腫瘍径 1 cm 前後の小型腺癌が多数発見される，③臨床病期 I 期例が約 70～80% を占める，等である．生存率については Sobue らの報告[1] しかないが，発見肺癌の予後が極めてよいことに関しても，ほぼ周知の事実であろう．

3. 癌検診の評価方法とバイアス

「CT 検診の発見率が胸部単純 X 線検診の数倍になる」という表現は，「胸部単純 X 線検診では放置すれば 1 年以内に顕在化する肺癌の大半を発見できてなくて，CT 検診ではそれらをすべて発見できる」という意味ではない．従来 1 年以内に顕在化する肺癌に対する胸部単純 X 線の感度は 71.6～75.0%[5,6] と報告されている．したがって「CT 検診の発見率が…数倍になる」という表現は，「放置すれば数年以内に顕在化するかもしれない肺癌を，CT 検診は 1 回で発見している」という意味である．一般に腫瘍倍加速度は時間を変数とした指数関数に沿って増大すると言われている．[7] Figure 1 に示すように，胸部単純 X 線で発見しうる腫瘍の大きさは腫瘍倍加曲線の傾きの強いところに相当し，CT でのみ発見しうる大きさは，逆にこの曲線の傾きのゆるやかな部分に相当すると考えられる．腫瘍倍加速度の傾きがゆるやかなほど，発見可能前臨床期 (preclinical detectable duration) は大きく延長する．定期検診を行った場合，発見可能前臨床期が長ければ，この期間内に検診を受診する確率が高くなるので，発見される確率も上昇する (length bias)．このように腫瘍倍加速度の遅いものほど発見率は向上しやすい．このように腫瘍倍加速度の遅い癌の術後生存率が極めて高いことから，検診の効果も高いと誤解しやすい．しかし，腫瘍倍加速度の遅いものは逆に放置しても顕在化するまで時間がかかるものであり，健在化しないままに他の病気で先に死亡するかもしれない．これを over-diagnosis bias と呼ぶ．以上のように発見された癌の特性を無視して，ただ予後のみによって検診の手法を評価することは，誤った解釈につながる．

癌検診の評価方法を，Table 2 にまとめた．現状では低線量 CT 検診には，症例研究までの成績しか報告されておらず，感度・特異度といった検診で発見できなかった癌を踏まえた評価さえも，いまだ報告されていない．

4. Japan Lung Cancer Screening Study の概要

我が国で開発された CT 検診の有効性評価に関しては，平成 11 年度に老人保健事業推進費等補助金「肺がん検診における高速らせん CT 法の効果評価研究」班が組織され，ランダム化比較試験を含んだ研究計画書が作成された．ランダム化比較試験の実現に向けて当時の厚生省老人保健課は相当なる努力を図ったものの予算上の問題からついに実現には至らなかった．その代わりとして，すでに行われた CT 検診受診者を追跡するコホート研究計画が実現化し，平成 13 年度に 21 世紀型医療開拓推進

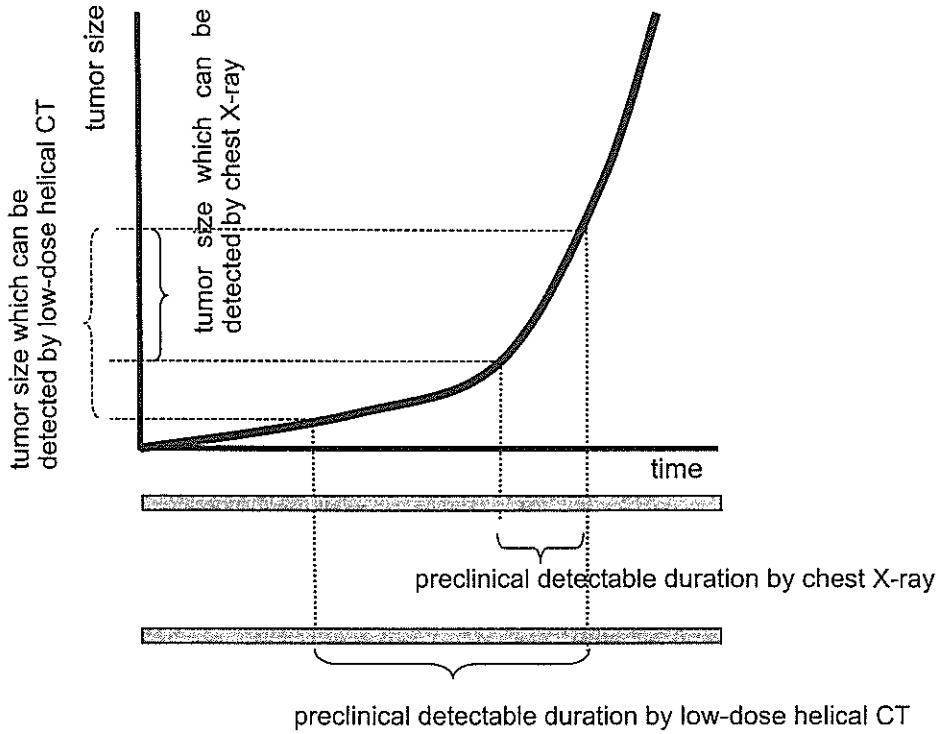

Figure 1. The curve of the relation between tumor size and time, comparing preclinical detectable duration by chest X-ray with that by low-dose helical CT.

Table 2. Study Design to Evaluate Cancer Screening

Method	Study design	Endpoint
Experimental study	Randomized controlled trial	Mortality reduction
Observational study	Cohort study	Mortality reduction
	Case-control study	Mortality reduction
		Sensitivity/Specificity
		Survival rate
	Case study	Resectability
		The rate of early stage
		Tumor size

費等補助金「がんの罹患高危険群の抽出と予後改善のための早期診断及び早期治療に関する研究」班が組織され，平成16年度からは第3次対がん総合戦略研究事業「革新的な診断技術を用いたこれからの肺がん検診手法の確立に関する研究」班と名を替え，研究を行っている．

研究デザインをFigure 2に示した．低線量CT検診を40歳以上で少なくとも一度受診したものを“CT検診群”とし，同じ時期に単純X線検診を受診しCT検診を以後受診しなかったものを“通常検診群”と定義した．追跡は主に住民基本台帳により異動を確認し，死亡者については，総務省からの許可を得た上で，人口動態調査死亡小票の閲覧を行い，死因を把握した．Table 3に各9地区の両群の登録者数を示す．CT検診群に46,733人，通常検診群に91,970人が登録されている．Table 4に2002年末までの追跡状況を元に得られた粗死亡率を示す．一見，肺癌死亡率は男女ともCT検診群の方が通常検診群よりも下回っているように見えるが，全死因に関しても同様の傾向が見られる．本研究は，ランダム化比較試験のように両群の性・年齢・喫煙を調整した研究ではなく，受診者をできるだけ制限なく登録した研究のため，両群の登録者の性・年齢・喫煙の分布には明らかな差がある．したがって粗死亡率の比較はあまり意味をなさず，今後層別化分析や多変量解析が必要となる．

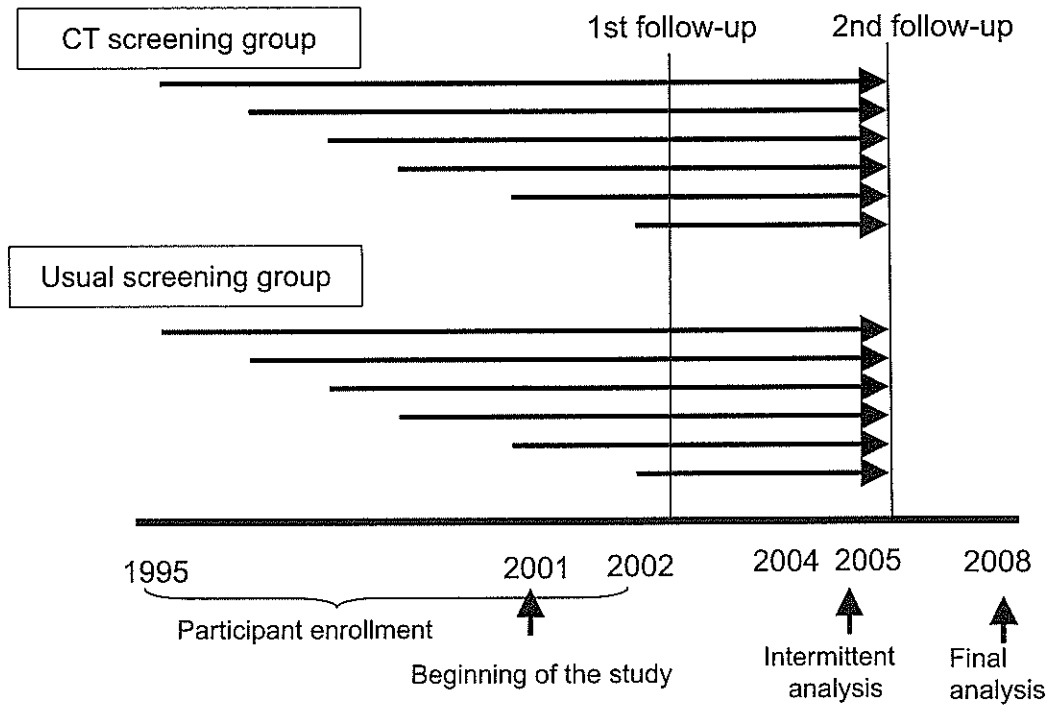

Figure 2. The Japan Lung Cancer Screening Study.

Table 3. Study Participants of Japan Lung Cancer Screening Study

	CT screening group		Usual screening group	
	Male	Female	Male	Female
Chiba	2,031	2,333	3,475	7,541
Tokyo	927	942	4,371	5,117
Hitachi	8,218	1,902	0	0
Niigata	5,306	1,323	7,972	4,147
Kanagawa	1,300	527	3,389	6,359
Osaka	2,766	1,925	4,181	9,201
Nagano	4,200	3,573	7,341	15,090
Okayama	827	57	1,168	122
Ehime	4,034	4,542	4,539	7,957
Total	29,609	17,124	36,436	55,534

Table 4. Crude Mortality of Japan Lung Screening Study

	CT screening group		Usual screening group	
	Male [104,055]	Female [59,078]	Male [179,246]	Female [283,881]
Lung cancer	76 (73.0)	10 (16.9)	180 (100.4)	61 (21.5)
All causes	683 (656.4)	163 (275.9)	2,103 (1173.2)	1,362 (479.8)

[ ]: follow-up (person-years). ( ): mortality per 100,000 person-years.

Table 5. Effective Doses at Chest X-ray Examination for Adult Male

Modality	Settings	Tube current (mA)	Effective dose (mSv)	
Miniature photofluorography	Screening	3.9	0.07	ref 12)
SDCT	Screening	50	1.40	ref 12)
SDCT	Clinical	100	2.74	ref 12)
MDCT (4-lows)	Clinical	127	10.02	ref 13)
MDCT (16-lows)	Clinical	175	9.36	ref 13)
MDCT (4-lows)	Screening	50	3.94	*
MDCT (16-lows)	Screening	50	2.74	*

SDCT: single-detector computed radiography, MDCT: multi-detector computed radiography.

*: These effective doses were estimated from the data of MDCT in the clinical setting based on the advice of Dr Nishizawa.

5. CT 検診に伴う不利益

CT 検診に伴う不利益については，開始当初より次の三つのことが問題視されてきた．

①過剰な要精検率

今までの報告例によると要精検率は 2～25% と報告されており，従来の単純 X 線撮影の 2～4% に比べてはるかに大きく，またバラツキが大きい．たとえ癌発見率が 0.3% (10 万対 300) であっても，要精検率が 10% であれば，要精検者 100 人中 97 人は癌ではなかったことになり，この 97 人に対して無駄な精密検査と精神的ダメージを与えることになる．検診はあくまで無症状者を対象にするもので，有症状で病院を訪れる患者とは異なった対応が必要であり，要精検率はできうる限り低いものでないと運用できない．

②高いコスト

従来の単純 X 線検診は，極めて安価であり，1 件あたり 1,000～1,500 円程度にすぎなかった．しかし CT は精密検査機器として開発されたものであり，高機能高価格なものが中心であるため，一般的には 1 件あたり平均 8,200 円程度で運用されている．[8] このような高額では検査の普及を図ることは困難であり，検診に特化した低機能低価格機種の開発が必要である．

③放射線被曝

一般に，日本人の年間平均自然放射線被曝は 2.4 mSv，年間平均医療放射線被曝は 2.25 mSv とされている．[9] 従来，放射線の健康影響は 100 mSv 未満では疫学的に確認されていないものの，国際的には 100 mSv 未満であっても影響があるという立場 (linear non-threshold theory：LNT 仮説) が採用されている．最近出された日本の医療放射線被曝に警鐘を投げかける二つの報告を紹介する．一つは 2004 年に Lancet に掲載されたイギリスと 14 カ国の医療放射線被曝を比較した分析である．この論文によれば日本人は X 線検査を年間平均 1,000 人対 1,477 件受けていると推定されており，日本人の癌死亡の 3.2% が医療放射線被曝によるものと推定している．[10] また 2005 年に British Medical Journal に掲載された 15 カ国原子力発電所従事者のコホート研究によれば，原子力発電所従事者 1 人あたりの累積平均被曝線量は 19.4 mSv で，白血病を除く全癌死亡について，1 Sv あたりの過剰相対リスクは 0.97 (95% 信頼区間：0.14～1.97) で，統計学的有意に死亡リスクが上昇したと報告されている．[11] この二つの論文は，方法論上にいくつかの大きな問題があり，その結果については懐疑的な意見も多いが，従来，医療用放射線被曝に対して寛容であった我が国の医療全体に冷や水をかけるようなものであった．さて，CT 検診の被曝はどうであろうか？ Table 5 に男性を対象とした実効線量を示す．[12,13]

従来用いられてきた間接撮影法は 0.07 mSv と非常に低い線量であるが，低線量 CT はシングルディテクターで 1.40 mSv，マルチディテクターで 2.74～3.94 mSv と推定されている．精密検査としての高分解能 CT はおそらくシングルディテクターで約 3 mSv，マルチディテクターで 10 mSv 以上と推定される．線量をどこまで軽減できるか，高分解能 CT による follow up をどこまで減らすことができるかが，CT 検診にとって極めて大きな課題である．

6. まとめ

低線量 CT 検診の有効性には，いまだ症例研究程度の evidence しか存在せず，無症状者を対象とした“検診”としての運用・普及は時期尚早と言わざるを得ない．癌検診の有効性評価として evidence level の高い感度・特異度や死亡率減少効果等については，今後の報告を待たざ