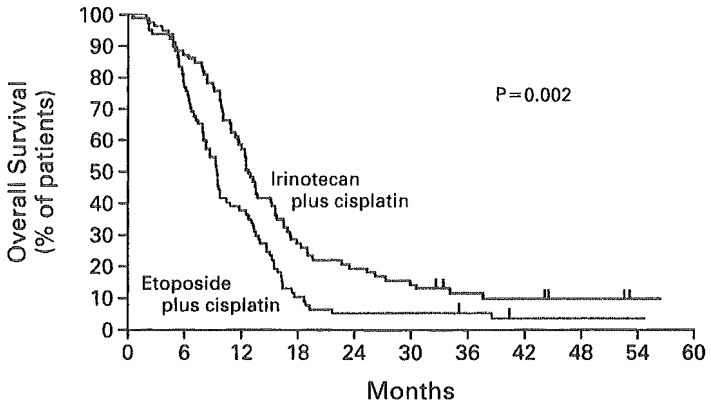
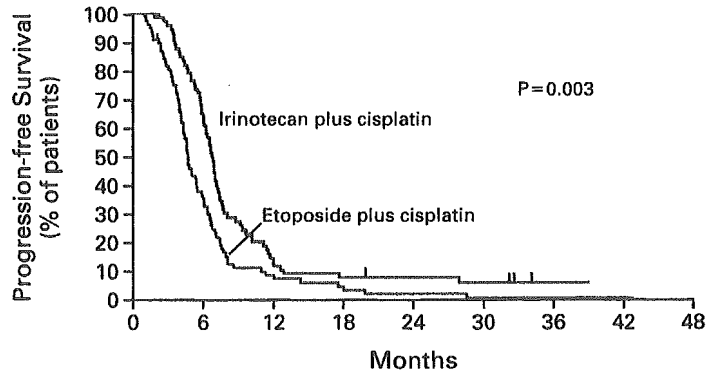


No. at Risk							
Irinotecan plus cisplatin	77	67	45	21	15	11	7
Etoposide plus cisplatin	77	60	29	8	4	4	3

Figure 1. Overall Survival of Patients with Extensive Small-Cell Lung Cancer Who Were Assigned to Treatment with Irinotecan plus Cisplatin or Etoposide plus Cisplatin.
The tick marks indicate patients whose data were censored.



No. at Risk				
Irinotecan plus cisplatin	77	47	9	9
Etoposide plus cisplatin	77	27	6	2

Figure 2. Progression-free Survival of Patients with Extensive Small-Cell Lung Cancer Who Were Assigned to Treatment with Irinotecan plus Cisplatin or Etoposide plus Cisplatin.
The tick marks indicate patients whose data were censored.

cycles of chemotherapy was similar in the two groups (approximately 70 percent), and thus the observed difference in survival is not thought to be attributable to a difference in the actual delivery of treatment.

Our study had several weaknesses. The planned second randomization to allow us to assess the benefit of subsequent thoracic radiotherapy was not completed; the planned quality-of-life study was not completed; and full information concerning treatment after disease progression was not available. The estimates of overall survival, however, should be highly reliable because, as of March 2001 (the final analysis), no patient had been lost to follow-up.

We consider that the trend toward a higher complete-response rate in the etoposide-plus-cisplatin group than in the irinotecan-plus-cisplatin group is due to chance. Although it is possible that these results occurred by chance, we believe that the decision to ter-

minate the trial early was based on generally accepted scientific and ethical principles and that, despite the small sample size, we can conclude that the combination of irinotecan and cisplatin is an attractive option for patients with metastatic small-cell lung cancer who have a good performance status.

Supported in part by grants-in-aid for cancer research and for the Second-Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health, Labor, and Welfare (Tokyo).

We are indebted to Ms. M. Imai and Ms. M. Niimi for data management; to Dr. K. Yoshimura for analyses during periodic interim monitoring; and to Dr. H. Fukuda for direction of the JCOG data center and oversight of management of the study.

APPENDIX

This study was coordinated by the Japan Clinical Oncology Group (M. Shimoyama, former chairperson) and was performed with the cooperation of the following institutions and investigators: National Dohoku Hospital, Hokkaido (T. Fujikane, K. Takahashi, and Y. Yamazaki); Hokkaido Keiaikai Minami-ichijo Hospital, Hokkaido (A. Fujita); Asahikawa Medical College Hospital, Hokkaido (Y. Osaki and Y. Nishizaki); Yamagata Prefectural Central Hospital, Yamagata (T. Tsukamoto); Tsukuba University Hospital, Ibaragi (S. Hasegawa and M. Tajima); Tochigi Cancer Center, Tochigi (T. Hirose, S. Machida, and M. Noda); National Nishi-Gunma Hospital, Gunma (S. Tsuchiya and H. Nakano); Saitama Cancer Center, Saitama (S. Yoneda, H. Sakai, T. Ikeda, and K. Kobayashi); National Cancer Center Hospital East, Chiba (F. Houjo, R. Kakinuma, Y. Ohe, T. Matsumoto, H. Ohmatsu, K. Kodama, E. Moriyama, and Y. Hosomi); National Cancer Center Central Hospital, Tokyo (T. Shinkai, H. Kunitoh, K. Kubota, and I. Sekine); International Medical Center of Japan, Tokyo (K. Kudo and Y. Takeda); Kanagawa Cancer Center, Yokohama (I. Nomura, K. Yamada, F. Oshita, Y. Kato, and M. Kondo); Yokohama Municipal Citizen's Hospital, Yokohama (H. Kunikane and A. Nagatomo); Niigata Cancer Center Hospital, Niigata (H. Tsukada, S. Mitsuma, and Y. Ichikawa); Aichi Cancer Center, Nagoya (K. Yoshida and T. Hida); National Nagoya Hospital, Nagoya (K. Nishiwaki and M. Hiraiwa); National Kinki Central Hospital for Chest Diseases, Osaka (M. Ogawara, T. Tsuchiyama, N. Kodama, K. Moriya, K. Okishio, N. Naka, S. Nobuyama, and S. Yamamoto); Kinki University School of Medicine, Osaka (N. Yamamoto, K. Nakagawa, T. Nogami, Y. Ieda, and M. Yoshida); Osaka Prefectural Habikino Hospital, Osaka (I. Kawase, N. Masuda, T. Nitta, and M. Kobayashi); Osaka City General Hospital, Osaka (K. Takeda, N. Yoshimura, H. Uejima, N. Nishikubo, T. Nitta, N. Takifuji, R. Miyaguchi, and K. Sugioka); National Toneyama Hospital for Chest Diseases, Osaka (H. Nishikawa and K. Shinkawa); Hyogo Medical Center for Adults, Hyogo (Y. Takada and T. Kadoh); Hyogo Medical College, Hyogo (K. Higashino and A. Tonomura); Wakayama Rohsai Hospital, Wakayama (T. Hoso and H. Minakada); Sasebo Municipal General Hospital, Nagasaki (J. Araki, K. Yamaguchi, and K. Ohba); Kumamoto Chuo Hospital, Kumamoto (T. Kiyama and Y. Yoshioka); and Kumamoto Regional General Hospital, Kumamoto (H. Senba and T. Seto).

REFERENCES

1. Fukuoka M, Furuse K, Saijo N, et al. Randomized trial of cyclophosphamide, doxorubicin, and vincristine versus cisplatin and etoposide versus alternation of these regimens in small-cell lung cancer. J Natl Cancer Inst 1991;83:855-61.
2. Roth BJ, Johnson DH, Einhorn LH, et al. Randomized study of cyclophosphamide, doxorubicin, and vincristine versus etoposide and cisplatin versus alternation of these two regimens in extensive small-cell lung cancer: a phase III trial of the Southeastern Cancer Study Group. J Clin Oncol 1992;10:282-91.
3. Ihde DC. Chemotherapy of lung cancer. N Engl J Med 1992;327:1434-41.
4. Aisner J. Extensive-disease small-cell lung cancer: the thrill of victory; the agony of defeat. J Clin Oncol 1996;14:658-65.
5. Masuda N, Fukuoka M, Kusunoki Y, et al. CPT-11: a new derivative of camptothecin for the treatment of refractory or relapsed small-cell lung cancer. J Clin Oncol 1992;10:1225-9.
6. Kudoh S, Fujiwara Y, Takada Y, et al. Phase II study of irinotecan combined with cisplatin in patients with previously untreated small-cell lung cancer. J Clin Oncol 1998;16:1068-74.
7. Tobinai K, Kohno A, Shimada Y, et al. Toxicity grading criteria of the Japan Clinical Oncology Group. Jpn J Clin Oncol 1993;23:250-7.
8. WHO handbook for reporting the results of cancer treatment. No. 48. Geneva: World Health Organization, 1979.
9. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993;85:365-76.
10. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958;53:457-81.
11. DeMets DL, Lan KKG. Interim analysis: the alpha spending function approach. Stat Med 1994;13:1341-52.
12. Reboussin DM, DeMets DL, Kim KM, Lan KKG. Programs for computing group sequential boundaries using the Lan-DeMets methods, version 2, 1998. (Accessed December 7, 2001, at http://www.biostat.wisc.edu/landemets/.)

Quality of Life Research **11**: 483–493, 2002.

The quality of life questionnaire for cancer patients treated with anticancer drugs (QOL-ACD): Validity and reliability in Japanese patients with advanced non-small-cell lung cancer

T. Matsumoto[1,2], Y. Ohashi[1,3], S. Morita[1,3], K. Kobayashi[1,4], M. Shibuya[1,5], Y. Yamaji[1,6], K. Eguchi[1,7], M. Fukuoka[8], K. Nagao[9], Y. Nishiwaki[2] & H. Niitani[10] for the CPT-11 Lung Cancer Study Groups West and East*

[1]*The QOL Evaluation Committee of the CPT-11 Phase III Lung Cancer Study Groups East and West;* [2]*National Cancer Center Hospital East, Chiba (E-mail: tsmatsum@nifty.com);* [3]*University of Tokyo, Tokyo;* [4]*Saitama Cancer Center, Saitama;* [5]*Nippon Medical School, Tokyo;* [6]*Mitoyo General Hospital, Kagawa;* [7]*National Shikoku Cancer Center Hospital, Ehime;* [8]*Kinki University, Osaka;* [9]*Chiba University, Chiba;* [10]*The Tokyo Cooperative Oncology Group, Tokyo, Japan*

Accepted in revised form 21 January 2002

Abstract

The quality of life questionnaire for cancer patients treated with anticancer drugs (QOL-ACD), which consists of four domains (functional, physical, mental, and psychosocial) and a global face scale, was developed as a generic questionnaire for Japanese cancer patients undergoing chemotherapy. We examined the validity and reliability of this questionnaire in Japanese patients with advanced non-small-cell lung cancer (NSCLC), who participated in two randomized phase III trials. After excluding two items, one showing low test–retest reliability and the other showing poor convergent validity for the target population, Cronbach's α coefficients ranged from 0.795 to 0.897 and the intra-class correlation coefficients ranged from 0.612 to 0.866. These results confirmed the high reliability of the questionnaire. The results of factor analysis provided strong support for the domain structure used in the questionnaire. Each of the four domains had a moderate to strong association with important clinical variables, such as performance status or weight loss, and correlation analysis showed that the face scale provided an appropriate measure of the global quality of life. These results indicated that the QOL-ACD is potentially useful for clinical research on Japanese patients with advanced NSCLC.

Key words: Item exclusion, Quality of life, QOL-ACD, Reliability, Validity

Abbreviations: ANOVA – analysis of variance; ECOG – Eastern Cooperative Oncology Group; EORTC QLQ-C30 – European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30; FACT-G – functional assessment of cancer therapy-general; NSCLC – non-small-cell lung cancer; PS – performance status; QOL – quality of life; QOL-ACD – quality of life questionnaire for cancer patients treated with anticancer drugs.

Introduction

Today, there are no researchers who deny the importance of quality of life (QOL) assessment for cancer patients, and the cancer patients themselves may also give priority to maintaining a high QOL rather than to survival when selecting treatment. In this respect, patients as well as researchers may have 'tacit knowledge', a term introduced by Polanyi as a key for the understanding of QOL [1].

*See Appendix B for details.

Test–retest reliability was evaluated in patients who filled out the questionnaire twice within 2–7 days. In addition, to evaluate the responsiveness of QOL assessment to changes of the health status over time, data obtained during the last week of the first course of chemotherapy were also used.

Statistical analyses

Our primary psychometric analysis to evaluate the validity and reliability of the QOL-ACD consisted of two steps: the first was the item-exclusion step, in which items were excluded if inappropriate in terms of psychometric relevance [19, 20], the second step was to confirm the domain structure of the questionnaire after excluding the inappropriate items. Furthermore, correlation analysis was performed to evaluate the appropriateness of the face scale as a measure of global QOL, and multivariate linear regression analysis was carried out to examine the clinical validity of the questionnaire.

In all analyses, scores for each of the four domains were calculated by summing the scores of items in each domain and then dividing the total by the number of items in the domain. When individual items were missing, the average of non-missing items within the same domain was used to replace the missing values. This rule was applied when at least half the items were completed in the domain. Otherwise, the score for the domain was treated as missing.

Item exclusion

To examine the psychometric properties of each item of the QOL-ACD, we assessed the test–retest reliability of the items using intra-class correlation coefficients, and we examined the convergent and discriminant validity using multi-trait scaling analysis. Multi-trait scaling analysis was performed using MAP-R software (Health Assessment Lab, Boston, MA) [21].

Confirmation of the domains and global QOL assessment

After the exclusion of inappropriate items, the internal consistency and test–retest reliability of each domain were evaluated using Cronbach's α coefficients and intra-class correlation coefficients, respectively. Factor analysis using promax rotation was also conducted to confirm the domain structure of the QOL-ACD. In addition, Pearson's correlation coefficients were calculated among the four domains to confirm the distinctiveness of each component of the QOL construct. It was hypothesized that conceptually related domains (e.g., functional and physical) would show a substantial correlation (Pearson's correlation coefficient above 0.40) with each other. It was considered undesirable for the correlation between any two domains was to be too high (e.g., above 0.70), because it would raise doubts about their distinctiveness [10, 16]. Pearson's correlation coefficients for the relationship of the face scale with the four domains were also calculated to evaluate the appropriateness of the face scale as a measure of global QOL. It was expected that the face scale would show a substantial correlation with each domain.

Clinical validity

Relationships between each domain and the baseline characteristics of the patients were examined using multi-variate linear regression analysis, where the baseline score for each domain and the baseline clinical characteristics were used as the response and explanatory variables, respectively. The baseline characteristics included PS (2/1/0), stage (IV/IIIB), sex (male/female), age ($\geqslant$65/$\leqslant$64 years), weight loss within the previous 6 months ($\geqslant$5%/$<$5%), albumin ($<$3.5/$\geqslant$3.5 g/dl) and LDH ($\geqslant$500/$<$500 IU/l), and they were categorized as shown in parentheses. Because it was expected that the psychosocial domain would be particularly related to the age and sex of the patients, two-way analysis of variance (ANOVA) was carried out to assess the relationship, including the interaction effect.

In order to evaluate the responsiveness of the questionnaire to changes in the health status, the relationship between changes of the score for each domain and the PS was analyzed from baseline to the last week of the first chemotherapy course using one-way ANOVA. Patients were categorized into three subgroups according to the change in PS, i.e., those whose PS improved, those whose PS was unchanged, and those whose PS deteriorated.

Apart from the multi-trait scaling analysis, all these analyses were performed using SAS software (SAS for Windows release 6.12, SAS Institute Inc., NC) [22–25].

Results

Patient characteristics

Of the 583 eligible patients in the two trials, 395 completed the pretreatment questionnaire at least once. When compared with the patients who did not fill out a pretreatment questionnaire, these patients had a slightly better PS and slightly less weight loss at baseline (Table 1). However, there were no large differences of patient characteristics between the two groups (Table 1).

Feasibility of using the questionnaire

Of the 583 eligible patients, 433 (74.3%) completed the questionnaire at least once during the clinical trial. Among the 150 patients whose questionnaires could not be collected, 71 lost the questionnaire, 47 were not handed the questionnaire form, 21 refused to fill in the questionnaire, and 11 did not fill it in for other reasons. These results suggested that data collection could be improved by better organization and more staff support. The percentage of missing answers to each item of the questionnaire is listed in Table 2. There was a low incidence of missing answers (range: 0.0–2.0%), except for item 6 (3.5%).

Table 1. Baseline characteristics of the study population

Characteristic	West N = 226	East N = 169	Total N = 395	Rest* N = 188
Age				
Median	64	62	63	63
Range	35–75	37–75	35–75	35–75
Albumin				
Median	3.8	3.9	3.8	3.7
Range	2.7–4.6	2.6–5.0	2.6–5.0	2.3–4.8
LDH				
Median	293	353	327	287.5
Range	83–1679	122–3246	83–3246	115–3107
Sex				
Male	168 (74.3)	130 (76.9)	298 (75.4)	148 (78.7)
Female	58 (25.7)	39 (23.1)	97 (24.6)	40 (21.3)
PS				
0	53 (23.5)	49 (29.0)	102 (25.8)	38 (20.2)
1	164 (72.6)	113 (66.9)	277 (70.1)	133 (70.7)
2	9 (3.9)	7 (4.1)	16 (4.1)	17 (9.1)
Stage				
IIIB	83 (36.7)	72 (42.6)	155 (39.2)	70 (37.2)
IV	143 (63.3)	97 (57.4)	240 (60.8)	118 (62.8)
Weight loss (%)				
⩾5	43 (19.0)	39 (23.1)	82 (20.8)	50 (26.6)
<5	160 (70.8)	111 (65.7)	271 (68.6)	113 (60.1)
Unknown	23 (10.2)	19 (11.2)	42 (10.6)	25 (13.3)

* Patients who did not complete a questionnaire at baseline, include 38 patients who only filled in the questionnaire during treatment.
The percentage are given in paranthesis.

Item exclusion

With regard to test–retest reliability, intra-class correlation coefficients showed that only item 6 had a substantially lower reliability (Table 2). Multi-trait scaling analysis showed that item–domain correlation coefficients were greater than 0.40, except for items 10 and 16 ($r = 0.37$ and 0.35, respectively). Accordingly, we used 0.40 as the criterion for item-convergent validity [10, 16]. In terms of item-discriminant validity, no definitive

Table 2. Percentage of missing answers and test–retest reliability of each item

Item number	Answers missing (%) N = 395	ICC[a] N = 107
1	0.3	0.627
2	2.0	0.483
3	1.5	0.501
4	0.5	0.603
5	1.5	0.612
6	3.5	0.217
7	0.0	0.563
8	0.5	0.553
9	0.3	0.615
10	0.5	0.387
11	0.8	0.753
12	0.5	0.686
13	0.5	0.606
14	0.0	0.670
15	0.3	0.731
16	0.8	0.668
17	1.5	0.715
18	0.3	0.758
19	0.8	0.832
20	0.0	0.811
21	0.3	0.863
22	0.3	0.751

[a] Intra-class correlation coefficient.

scaling error occurred. The finding that item 10 was not clearly included in any domain and had a low reliability (intra-class correlation 0.387) could be due to the fact that the scores for this item were highly skewed (because only a few patients reported mild vomiting at baseline). Since vomiting is one of the major adverse events induced by chemotherapy, we thought that it was still necessary to include item 10 in the QOL assessment despite this result. However, we judged that items 6 and 16 should be excluded because of low test–retest reliability and poor convergent validity, respectively.

Confirmation of the domain structure and global QOL assessment

After exclusion of items 6 and 16, the uniformly high Cronbach's α coefficients (0.795–0.897) as well as intra-class correlation coefficients (0.610–0.866) of the domain scores confirmed the internal consistency and test–retest reliability of the questionnaire (Table 3). Factor analysis provided strong support for the domain structure, as shown by the pattern of factor loadings in Table 4. Table 5 presents Pearson's correlation coefficients for the relations among the four domains and for the relationship of each domain with the face scale. All correlations among the four domains were lower than 0.70. As expected, the functional and physical domains showed a relatively good correlation with each other. Modest correlations (Pearson correlation coefficient >0.40) were found between the functional and mental domains ($r = 0.618$), physical and mental domains ($r = 0.588$), and mental and psychosocial domains ($r = 0.517$). The face scale was most strongly correlated with the mental domain, while it showed substantial correlations with the physical and psychosocial domains and a weaker relationship with the functional domain.

Clinical validity

The results of multivariate linear regression analysis to evaluate the relationship between each domain and baseline clinical characteristics are shown in Table 6. Statistically significant relationships were found between the functional domain and PS, sex, or weight loss; between the physical domain and PS, weight loss, albumin, or LDH; between the mental domain and PS or weight loss; and between the psychosocial domain and sex or age. Two-way ANOVA showed that younger women had more deterioration of the psychosocial domain (sex: $p = 0.005$, age: $p < 0.001$, the interaction between sex and age: $p = 0.038$).

One-way ANOVA was used to evaluate the sensitivity of the questionnaire to changes in PS over time and the results are shown in Table 7. A significant relationship was found between changes of PS and changes of the functional, physical, and mental domains ($p = 0.015$, 0.005, and 0.008, respectively).

Discussion

In the late 1980s, development of QOL questionnaires was started in Japan by two approaches [12]. First, well-established QOL questionnaires developed in English-speaking or western European countries, such as the FLIC, EORTC QLQ-C30 and FACT-G, were translated into Japanese. Second, a number of new QOL questionnaires were developed by taking the Japanese perspective of QOL into account. In this way, the QOL-ACD was developed by the Japanese Quality of Life Research Group (comprising doctors, nurses, psychologists, and patients).

Eguchi et al. [26] translated Schipper's FLIC into Japanese and studied its reliability and va-

Table 3. Descriptive statistics, internal consistency and test–retest reliability of the QOL-ACD domains (n = 394)

Domain	No. of items	Mean score	SD	ICC[a]	Cronbach's α
Functional	5	4.12	0.96	0.610	0.897
Physical	5	4.14	0.70	0.708	0.795
Mental	4	3.64	0.82	0.788	0.809
Psychosocial	5	3.15	0.95	0.866	0.818

[a] Intra-class correlation coefficient (n = 107).

Appendix B

The following institutions participated in this study:

Group West: Osaka City General Hospital, Osaka; Osaka Prefectural Habikino Hospital, Osaka; Hyogo Medical Center for Adults, Hyogo; Aichi Cancer Center, Aichi; Osaka City University Medical School, Osaka; Kobe City General Hospital, Hyogo; Aichi Prefectural Hospital, Aichi; Osaka Medical Center for Cancer and Cardiovascular Diseases, Osaka; National Kinki Central Hospital for Chest Diseases, Osaka; Toneyama National Hospital, Osaka; Hyogo College of Medicine, Hyogo; Faculty of Medicine, Kyushu University, Fukuoka; Kumamoto Regional Medical Care Center, Kumamoto; Shimonoseki Kosei Hospital, Yamaguchi; Gifu Municipal Hospital, Gifu; Saiseikai Nakatsu Hospital, Osaka; Izumisano Municipal Hospital, Osaka; Osaka Prefectural General Hospital, Osaka; Japanese Red Cross Nagoya First Hospital, Aichi; Nagasaki University, School of Medicine, Nagasaki; Osaka Posts and Telecommunications Hospital, Osaka; Japanese Red Cross Society Wakayama Medical Center, Wakayama; Kagawa Prefectural Center Hospital, Kagawa; Saga Medical School, Saga; Sasebo City General Hospital, Nagasaki; School of Medicine, Kanazawa University, Ishikawa; Wakayama Rosai Hospital, Wakayama; School of Medicine, The University of Tokushima, Tokushima; Hiroshima University School of Medicine, Hiroshima; Chugoku Rosai Hospital, Hiroshima; National Minamifukuoka Chest Hospital, Fukuoka; Aso Cement Iizuka Hospital, Fukuoka; Kurume University, School of Medicine, Fukuoka; Kumamoto Chuo Hospital, Kumamoto; Nagasaki Municipal Citizens Hospital, Nagasaki; Gifu University School of Medicine, Gifu; Kobe University School of Medicine, Hyogo; Hiroshima Red Cross Hospital, Hiroshima; National Kyushu Cancer Center, Fukuoka; Mitoyo General Hospital, Kagawa.

Group East: National Cancer Center Hospital East, Chiba; National Cancer Center Hospital, Tokyo; Niigata Cancer Center Hospital, Niigata; Saitama Cancer Center, Saitama; Kanagawa Cancer Center, Kanagawa; National Nishi-Gunma Hospital, Gunma; School of Medicine, Chiba University, Chiba; Tochigi Cancer Center, Tochigi; National Shikoku Cancer Center, Ehime; Kanagawa Prefectural Cardiovascular and Respiratory Center, Kanagawa; National Dohoku Hospital; Tokyo Metropolitan Fuchu Hospital, Tokyo; National Tokyo Hospital, Tokyo; Nippon Medical School, Tokyo; Sendai Kousei Hospital, Sendai; Funabashi Municipal Medical Center, Chiba; Yokkaichi Municipal Hospital, Mie; School of Medicine, Gunma University, Gunma; Jichi Medical School, Tochigi; Sapporo National Hospital, Hokkaido; Okayama University Medical School, Okayama; National Okinawa Hospital, Okinawa; Yokohama Municipal Citizen's Hospital, Kanagawa; Asahikawa Medical College, Hokkaido; Nippon Medical School Chiba Hokusoh Hospital, Chiba; Okayama Red Cross Hospital, Okayama; Ehime Prefectural Central Hospital, Ehime; Okayama Rosai Hospital, Okayama; Sapporo Medical University, Hokkaido; Showa University School of Medicine, Tokyo.

References

1. Browman GP. Science, language, intuition, and the many meanings of quality of life. J Clin Oncol 1999; 17: 1651–1653.
2. Aaronson NK, Bullinger M, Ahmedzai S. A modular approach to quality-of-life assessment in cancer clinical trials. Recent Results Cancer Res 1988; 111: 231–249.
3. Machin D. Assessment of quality of life in clinical trials of the British Medical Research Council. J Natl Cancer Inst Monogr 1996; 20: 97–102.
4. Moinpour CM, Feigl P, Metch B, et al. Quality of life end points in cancer clinical trials: Reviews and recommendations. J Natl Cancer Inst 1989; 81: 485–495.
5. Nayfield SG, Ganz PA, Moinpour CM, et al. Report from a National Cancer Institute (USA) Workshop on quality of life assessment in cancer clinical trials. Qual Life Res 1992; 1: 203–210.
6. Cella DF, Tulsky DS. Quality of life in cancer: Definition, purpose, and method of measurement. Cancer Invest 1993; 11: 327–336.
7. Osoba D. Lessons learned from measuring health-related quality of life in oncology. J Clin Oncol 1994; 12: 608–616.
8. Rowland JH. Outcome assessment: Cancer-specific quality-of-life measures – beyond the research setting. Educational Book, ASCO 33rd Annual Meeting, Denver, CO, May 17–20, 1997: 342–349.
9. Schipper H, Clinch J, McMurray A, et al. Measuring the quality of life of cancer patients: The Functional Living Index-Cancer: Development and validation. J Clin Oncol 1984; 2: 472–483.
10. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993; 85: 365–376.
11. Cella DF, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy Scale: Development and validation of the general measure. J Clin Oncol 1993; 11: 570–579.
12. Kobayashi K. Development of QOL questionnaires in Japan. Gan To Kagaku Ryoho 1999; 2: 472–483 (in Japanese).
13. Kobayashi K, Takeda F, Teramukai S, et al. A cross-validation of the European Organization for Research and Treatment of Cancer QLQ-C30 (EORTC QLQ-C30) for Japanese with lung cancer (see comments). Eur J Cancer 1998; 34: 810–815.
14. Fumimoto H, Kobayashi K, Chang CH, et al. Cross-cultural validation of an international questionnaire, the general measure of the Functional Assessment of Cancer Therapy (FACT-G), for Japanese. Qual Life Res 2001; 10: 201–209.
15. Kurihara M, Shimizu H, Tsuboi K, et al. Development of quality of life questionnaire in Japan: Quality of life assessment of cancer patients receiving chemotherapy. Psychooncology 1999; 8: 355–363.
16. Kaasa S, Bjordal K, Aaronson NK, et al. The EORTC Core Quality of Life Questionnaire (QLQ-C30): Validity and reliability when analyzed with patients treated with palliative radiotherapy. Eur J Cancer 1995; 31A(13/14): 2260–2263.
17. Masuda N, Fukuoka M, Negoro S, et al. Randomized trial comparing CDDP and CPT-11 versus CDDP and VDS versus CPT-11 alone in advanced NSCLC: A multicenter

phase III study. Proc Am Soc Clin Oncol 1999; 18: 459 (abstr).
18. Niho S, Nagao K, Nishiwaki Y, et al. Randomized multicenter phase III trial of irinotecan (CPT-11) and cisplatin (CDDP) versus CDDP and vindesine (VDS) in patients with advanced non-small cell lung cancer (NSCLC). Proc Am Soc Clin Oncol 1999; 18: 492 (abstr).
19. Stewart AL, Hays RD, Ware JE Jr. The MOS short-form general health survey. Reliability and validity in a patient population. Med Care 1988; 26: 724–735.
20. Ware JE Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992; 30: 473–483.
21. Ware JE, Harris W, Gandek B, et al. MAP-R for Windows: Multitrait/Multi-item Analysis Program – Revised User's Guide. Boston, MA: Health Assessment Lab, 1997.
22. SAS Institute, Inc. SAS Procedures Guide, Version 6, 3rd ed., SAS Institute, Inc. Cary, NC, 1990.
23. SAS Institute, Inc. SAS STAT User's Guide, Version 6, 4th ed., Vol. 1, SAS Institute, Inc. Cary, NC, 1990.
24. SAS Institute, Inc. SAS STAT User's Guide, Version 6, 4th ed., Vol. 2, SAS Institute, Inc. Cary, NC, 1990.
25. SAS Institute, Inc. SAS STAT Software: Changes and Enhancements through Release 6.12, SAS Institute, Inc. Cary, NC, 1997.
26. Eguchi K, Fukutani M, Kanazawa M, et al. Feasibility study on quality-of-life questionnaires for patients with advanced lung cancer. Jpn J Clin Oncol 1992; 22: 185–193.
27. Aaronson NK. Assessing the quality of life of patients with cancer: East meets West. Eur J Cancer 1998; 34: 767–769.
28. Nagata C, Ido M, Shimizu H. Choice of response scale for health measurement: comparison of 4, 5, and 7-point scales and visual analog scale. J Epidemiol 1996; 6: 192–197.
29. Ganz PA, Hirji K, Sim M, et al. Predicting psychosocial risk in patients with breast cancer. Med Care 1993; 31: 419–431.

Address for correspondence: Taketoshi Matsumoto, Masashigaoka Hospital, 7-15-1, Kusunoki, Kumamoto-shi, Kumamoto, 862-8003, Japan
E-mail: tsmatsum@nifty.com

Clinical Investigations

Respiration

Respiration 2002;69:30–37

Received: October 9, 2000
Accepted after revision: July 24, 2001

Correlation among the Indices of High-Resolution Computed Tomography, Pulmonary Function Tests, Pulmonary Perfusion Scans and Exercise Tolerance in Cases of Chronic Pulmonary Emphysema

Etsuo Fujita[a] Yukio Nagasaka[b] Takenorii Kozuka[c] Hidemitsu Ebara[d] Masahiro Fukuoka[e]

[a]National Kinki-Chuo Hospital for Chest Disease, Nagasone-cho, Sakai, Osaka, [b]Department of Respiratory Medicine, Sakai Hospital, Kinki University School of Medicine, Sakai, Osaka, [c]Department of Radiology, Osaka University School of Medicine, Osaka, [d]Department of Radiology, Matsuyama Municipal Hospital, Matsuyama, Ehime, [e]Fourth Department of Internal Medicine, Kinki University School of Medicine, Osakasayama, Osaka, Japan

Key Words
Emphysema · Treadmill · High-resolution computed tomography · Single photon emission computed tomography · $\dot{V}O_2$ peak

Abstract
Background: Mismatched distribution of pulmonary blood flow is a common characteristic in emphysematous patients. But few reports have mentioned the relationships between the morphological changes in the lungs as assessed by high-resolution computed tomography (HRCT), pulmonary blood flow (PBF) scan and the indices of exercise tolerance. We investigated these relationships. ***Objective:*** Pulmonary function tests (PFT), HRCT, single photon emission computed tomography (^{99m}SPECT) and treadmill exercise tests were performed on emphysematous patients, and the correlations between these examinations were studied. ***Methods:*** We evaluated 20 patients (M 18, F 2, age 66 ± 8.0 years). CT evaluation was performed according to the grade of emphysematous change. ^{99m}SPECT was performed to evaluate mismatched PBF by the score method. The better flow of the middle lobe was selected to be the standard lobe for the basic PBF. That score was set to 1 when the blood flow was below 60 or above 140%. PBF between 60 and 140% was scored as 0. ***Results:*** FEV_1 ($r = 0.648$, $p = 0.002$) and VC ($r = 0.767$, $p = 0.001$) correlated significantly with $\dot{V}O_2$ peak . FEV_1 ($r = 0.667$, $p = 0.0018$) correlated significantly with anaerobic threshold (AT). CT grade did not correlate with PBF mismatch score ($r = 0.266$, $p = 0.3376$). $\%\dot{V}O_2$ peak did not correlate with CT grade ($r = -0.467$, $p = 0.0689$) or with mismatch PBF score ($r = -0.327$, $p = 0.2377$). ***Conclusions:*** HRCT and ^{99m}SPECT were advantageous for detecting the progression of disease and emphysematous changes. However, the severity of anatomical emphysematous changes did not necessarily correlate with the indices of exercise tolerance and pulmonary function tests.

Introduction

High-resolution computed tomography (HRCT) is the best modality for detecting pulmonary emphysema [1]. There have been a few reports that the severity of emphysematous changes was correlated with exercise tolerance

KARGER
Fax + 41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0025–7931/02/0691–0030$18.50/0

Accessible online at:
www.karger.com/journals/res

Etsuo Fujita
National Kinki-Chuo Hospital for Chest Disease
1180 Nagasone-cho, Sakai
Osaka 591-855 (Japan)
Tel. +81 722 52 3021, Fax +81 722 51 1372, E-Mail efujita@kinchu.hosp.go.jp

Table 1. Pulmonary function tests

Name	Age Sex	HT cm	BW kg	Smoking history pack-years	VC	%VC %	FEV_1 liters	$\%FEV_1$ %	$FEV_1\%$ %	TLC liters	%TLC %	FRC liters	RV/TLC %	DLco ml/min mm Hg	%DLco %	$DLCO/V_A$ ml/min/ mm Hg/l	Increase of FEV_1 after orciprenaline	
																	liters	%incr.
HA	72 M	163	61	NA	3.18	116	1.59	59.6	43.1	6.69	114.1	4.1	44.8	13.75	89.8	2.64	NA	NA
SS	68 M	165	43	43	2.51	76.3	0.70	24.7	27.9	7.52	117.6	6.12	66.5	NA	NA	NA	NA	NA
YS	62 M	164	53	60	2.48	74.0	0.52	17.7	21.7	6.65	112.5	5.18	59	20.23	126.2	3.86	NA	NA
IH	68 M	160	58	NA	3.09	96.6	2.10	79.2	68.0	5.42	93.7	3.04	51.1	NA	NA	NA	NA	NA
TH	73 F	145	48	51	1.83	88.4	0.68	20.1	41.0	3.58	131.7	2.55	48.9	NA	NA	NA	NA	NA
NH	53 M	176	57	28	2.89	76.7	1.31	36.7	45.3	7.7	112.5	5.87	56.8	13.95	70.8	2.33	6.1	0.1
MT	62 M	170	60	84	4.62	131	2.41	76.5	53.2	8.4	130.4	6.13	44.5	8.86	48.4	1.41	NA	NA
DY	57 M	175	59	80	3.67	98.7	1.51	43.8	45	6.3	93.1	3.99	35.7	9.82	49.8	2.09	0	0
AT	71 M	167	54	Ex	3.18	97.2	0.78	27.5	33.4	6.18	100.1	4.41	45.2	6.33	43.8	1.34	NA	NA
KK	64 M	176	69	Ex	4.25	118	1.17	35.3	39.2	6.97	101	4.78	45.2	26.48	134	4.9	NA	NA
IM	60 M	156	43	31.5	2.47	75.8	0.69	25.6	27.9	5.52	101.9	4.36	46.4	7.04	44.3	1.83	NA	NA
SI	85 M	172	57	0	3.36	108	1.20	44.6	41.8	7.36	111.1	5.4	56.8	15.62	137.2	2.66	21.6	0.27
YJ	62 M	167	64	102	2.91	84.3	1.26	41.4	43.9	5.77	95.2	4.39	49.7	11	60.2	2.51	0.8	0.01
IH	70 M	156	48	69	3.21	104	1.10	44.7	34.2	6.82	128.9	4.95	50.6	11.26	88.3	2.08	13.1	0.1
FK	53 M	151	70	15	3.19	91.6	1.25	46.7	41.5	7.02	141.6	5.26	54.6	NA	NA	NA	NA	NA
TH	63 M	172	52	NA	3.96	112	2.33	72.9	60.8	NA	NA	NA	NA	NA	NA	NA	NA	NA
KO	77 F	147	36	120	1.59	77.1	0.88	25.6	55.7	NA	NA	NA	NA	NA	NA	NA	NA	NA
YM	74 M	160	74	25	2.97	96.1	1.22	48.6	41.0	7.34	126.9	5.5	60.4	9.44	88.2	1.64	2.3	0.03
HS	81 M	167	60	NA	1.97	63.5	0.96	36.9	49.7	4.65	80	3.24	53.3	11.73	91.4	3.26	0	0
MT	63 M	172	67	60	3.83	108	2.54	79.4	72.9	6.68	101.9	3.95	43.1	17.51	90.4	3.35	NA	NA
Mean	66.9	164.1	56.6	54.9	3.1	94.7	1.31	44.4	44.4	6.48	110.8	4.62	50.7	13.07	83.1	2.56	6.27	0.073
SD	8.6	9.3	9.9	32.9	0.8	17.8	0.61	19.8	13	1.13	16.0	1.01	7.3	5.29	31.2	0.96	8.23	0.09

NA = Not available, Ex = ex smoker.

[2]. We have already reported that the peak expiratory flow (PEF) index is very useful with asthmatic patients, not only in detecting aggravation of asthma but also in evaluating exercise tolerance [3, 4]. Emphysema is commonly defined as enlargement of the air spaces distal to the terminal bronchioles. The pathogenesis of emphysema is still unknown: the most accepted hypothesis is based on an imbalance between proteases and antiproteases [5].

The percent low-attenuation area (LAA%) is understood to be the most accurate index in quantifying emphysema by CT [6].

In a comparison of asthmatic and emphysematous patients matched for age and pulmonary function, the asthmatic patients had a better exercise capacity than the emphysematous patients [7]. Therefore, HRCT may be a good modality for detecting emphysematous change in asthmatic patients. Asthmatic patients have a CT density between that of a normal person and that of an emphysematous person [8]. Clinically, a 6-min walk distance can be easily performed on a treadmill, and performance on this test is sometimes used to measure exercise tolerance [9]. We quantitatively evaluated the exercise tolerance of emphysematous patients by using the treadmill, that is, as the emphysematous changes increased, exercise tolerance was reported to decrease [10].

Regarding the distribution of pulmonary blood flow (PBF), there were regional decreases or defects in PBF in emphysematous patients. Pulmonary perfusion scans are also useful for detecting the grade of mismatched distribution of PBF [11]. There have been few reports that mentioned the relationships between the morphological changes of lungs as determined by HRCT, the scores of mismatched PBF determined by single photon emission computed tomography (^{99m}SPECT) and exercise tolerance. In this study, we investigated the correlations between these examinations.

Patients and Methods

Patients

Twenty patients (male 18, female 2, age 66.9 ± 8.6 years) with chronic stable emphysema were studied (table 1). We obtained informed consent for HRCT, ^{99m}SPECT, pulmonary function tests (PFT) and exercise testing from each subject. Exercise tests were performed according to a previously described method [3, 4], and $\dot{V}O_2$

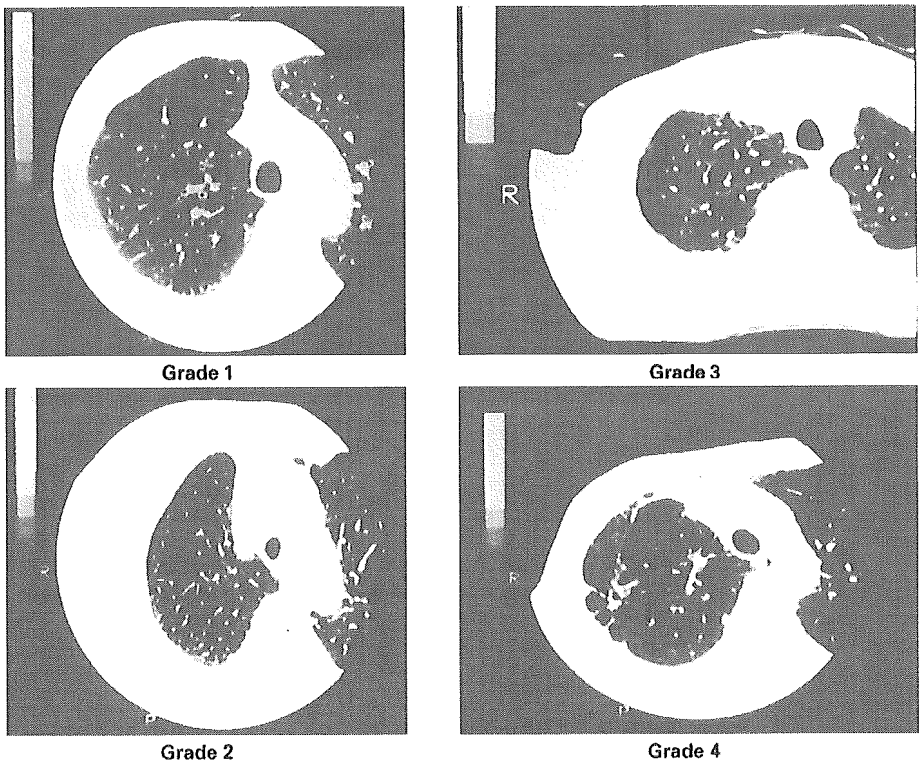


Fig. 1. Grading of CT in emphysematous patients. Emphysematous change became severe from grade 1 to 4. Gradings were divided by the distribution of LAA and their sizes. This grading was done according to a previously described method [12].

peak, carbon dioxide production ($\dot{V}CO_2$) peak, anaerobic threshold point (AT), heart rate (HR) at peak exercise and oxygen saturation (SaO_2) at peak exercise were measured.

CT and ^{99m}SPECT

HRCT was performed by CT scan (Type TSX, X Vigor, Toshiba, Japan). CT scans were performed with a 1-second scan time (120 kV, 250 mA), at 2-mm intervals from the apex of the lung to the diaphragm using 2 mm collision. The borders among upper, middle and lower parts of the lungs were set at the aortic arch and influx of the right lower pulmonary vein. Window width was 1,600 and window level was −650 IU. ^{99m}SPECT was performed on an SNC-510R-20 (Shimazu, Japan).

Lung Function Tests

One hour before exercise testing, LFT were performed and VC, FEV_1, $\%FEV_1$ and $FEV_1\%$ were measured with a spirometer (AS-7®, Minato Medical Co., Osaka, Japan). Diffusing capacity of the lung (DL_{CO}), total lung capacity (TLC), functional residual capacity (RC) and residual volume/TLC (RV/TLC) were measured by a total respiratory function automated analysis system (Fudac 50®, Fukuda Densi, Osaka, Japan). Bronchial hypersensitivity was checked beforehand by inhalation of orciprenaline (0.5 ml).

CT Grading

CT grading was evaluated in accordance with previously reported methods [12–14], with the grade of emphysematous change equaling one of four grades decided by the distribution and size of the LAA. The LAA was considered to be small (<5 mm), medium (>5 mm, <10 mm) or large (>10 mm). Taking the distribution into consideration, the LAA was sparse and small in grade 1, predominantly in the upper lobe and from small to medium in grade 2, in the upper and lower lobe and from medium to large in grade 3, in the upper and lower lobe and large in grade 4.

We adopted a grade of 0.5 when the emphysematous change was between 0 and 1. Six observers independently scored each CT scan using this grading system, and the final scores were determined by the majority vote (fig. 1).

^{99m}SPECT

A ^{99m}SPECT pulmonary perfusion scan was done in 14 patients, on the three parts of both lungs, labeled the upper, middle and lower parts (table 2), and we calculated the blood flow from the middle lobe to the upper, middle and lower parts. If the blood flow was below 60% or more than 140%, we assigned it a score of 1, and if it was between 60 and 140%, we assigned it a score of 0, following a previously reported method [15]. The better flow of the middle lobe was selected to be the standard lobe of basic PBF, and the percentage of each lobe in comparison to the standard middle lobe was calculated and a score of 0 or 1 was obtained for each of the 14 patients. The score for both lobes was summed. The scans and measurements obtained for case M.T. are shown in table 2 and figure 2 as an illustration. We measured the mismatched PBF score in the 14 patients.

Treadmill Exercise Tests

Exercise testing was performed by following a protocol of continuous incremental multiple loads. In that protocol, after a 1-min rest,

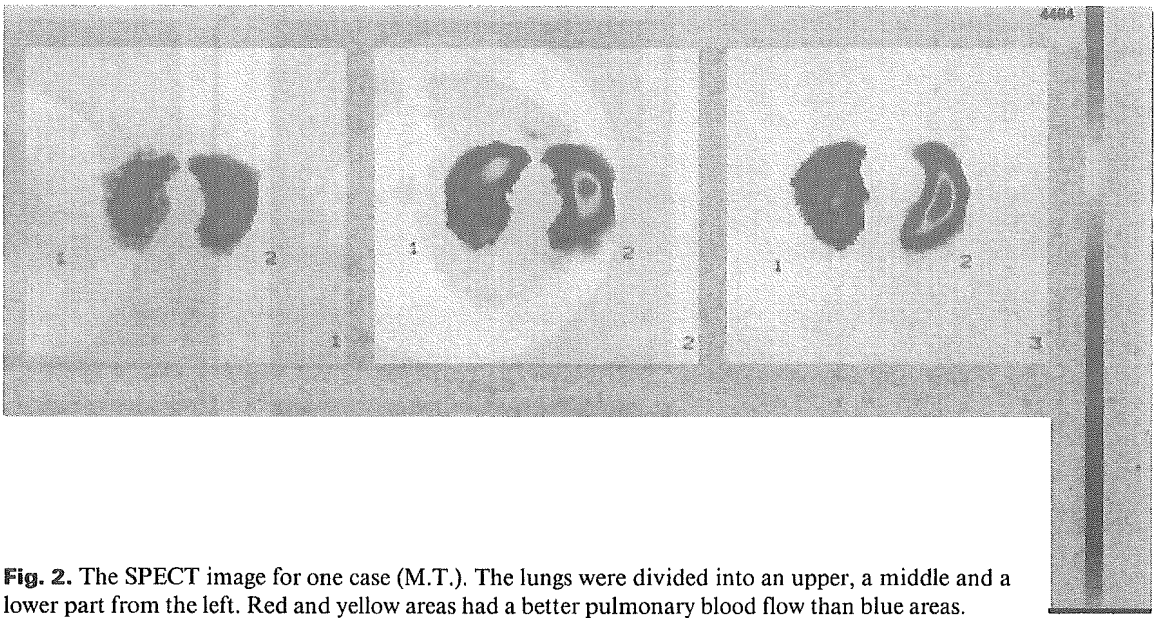


Fig. 2. The SPECT image for one case (M.T.). The lungs were divided into an upper, a middle and a lower part from the left. Red and yellow areas had a better pulmonary blood flow than blue areas.

Table 2. ^{99m}SPECT (Case M.T.)

Part	Scintillation count		Ratio of the base to middle lung		Mismatch score of pulmonary perfusion scan		
	left	right	left (%)	right (%)	left	right	sum
Upper lung	1,056	1,090	55.8	57.6	1	1	2
Middle lung	1,894	1,550	100	81.8	0	0	0
Lower lung	1,900	1,401	100.3	74.0	0	0	0
Total lung (score)							2

warming up was done for 4 min at 1 km/h. Slope was started at 8% and increased gradually by 1–3 or 4% at each 1-min interval [7]. The end point of the exercise test was at the occurrence of severe dyspnea (Borg scale >5) or fatigue, tachycardia (>80% of target HR) or desaturation (SaO_2 <90%). Target HR was determined as 220 – age (years)/min [3, 4]. The exercise was stopped at 80% of target HR [3]. Expired gases were continuously analyzed with an RM-300® (Minato Medical Co.). Minute ventilation ($\dot{V}E$), oxygen consumption ($\dot{V}O_2$), carbon dioxide production ($\dot{V}CO_2$), tidal volume (VT) and respiratory frequency (RR) were measured. SaO_2 and HR were monitored by pulse oxymetry and electrocardiogram (ECG). Severity of the sense of dyspnea was rated according to the Borg scale in response to the manual pushing of a button by patients. The AT was determined by the $\dot{V}$ slope method using the least squares method [16]. $\dot{V}O_2$ peak, $\dot{V}CO_2$ peak and $\dot{V}_E$ peak at AT ($\dot{V}O_2$ AT, $\dot{V}CO_2$ AT, $\dot{V}_E$ AT), SaO_2 at peak exercise, HR and SaO_2 at AT were also measured. The predicted value of $\dot{V}O_2$ was used from the study by Jones [17]. For males this was 60 – 0.55 × age (ml/kg/min), and for females this was 48 – 0.37 × age (ml/kg/min).

Statistics

The correlations between PFT were checked by simple regression methods and judged to be significant when $p > 0.05$ (fig. 3–5).

The correlations between indices of exercise tolerance, CT grading scores and the mismatch scores of PBF were examined by Sperman's rank correlation (table 4).

Results

Pulmonary Function Test

The results of PFT showed that VC was 3.1 ± 0.8 liters (mean ± SD), %VC was 94.7 ± 17.8%, FEV_1 was 1.31 ± 0.61 liters and $\%FEV_1$ was 44.4 ± 19.8%.

$FEV_1\%$ was 44.4 ± 13% (table 2). $\%DL_{CO}$ was 83.1 ± 31.2% (measurable in 14 patients). TLC was 6.48 ± 1.13 liters, FRC was 4.62 ± 1.01 liters and RV/TLC was 50.7

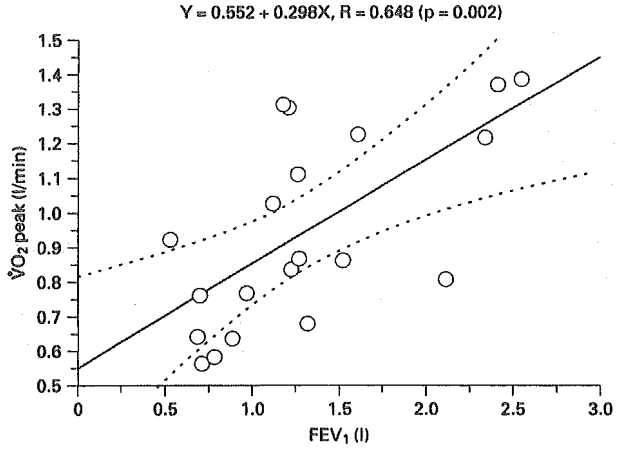


Fig. 3. FEV_1 correlated with $\dot{V}O_2$ peak (p = 0.002). Patients with higher FEV_1 showed a higher $\dot{V}O_2$ peak. The figure shows the 95% confidence bands for the true mean of the $\dot{V}O_2$ peak.

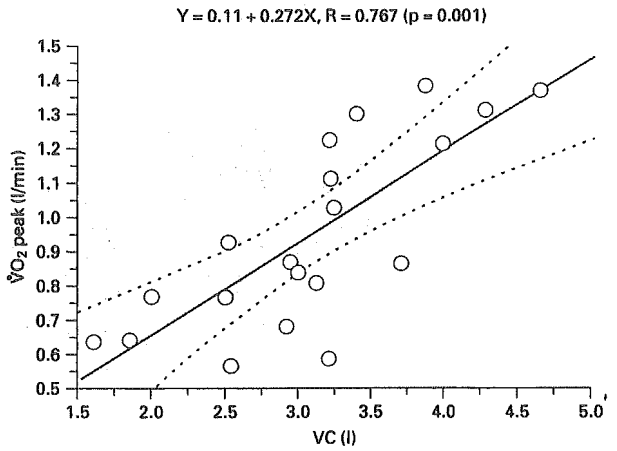


Fig. 4. VC correlated with $\dot{V}O_2$ peak (p = 0.001). Patients with higher VC showed a higher $\dot{V}O_2$ peak. The figure shows the 95% confidence bands for the true mean of the $\dot{V}O_2$ peak.

± 7.3% (measurable in 15 patients). Increase of FEV_1% after orciprenalline was 0.07 ± 0.09% (measurable in 7 patients) (table 1).

Outcome of Treadmill Exercise Tests

The results of treadmill exercise tests showed that $\dot{V}O_2$ peak was 0.928 ± 0.278 (l/min) and %$\dot{V}O_2$ peak was 74.06 ± 26.40%. $\dot{V}CO_2$ peak was 0.914 ± 0.295 (l/min). $\dot{V}_E$ peak was 37.46 ± 13.57. $\dot{V}O_2$ AT was 0.724 ± 0.222. $\dot{V}CO_2$ AT was 0.658 ± 0.199 (l/min). $\dot{V}_E$ at AT was 28.38 ± 7.16. SaO_2 at peak exercise was 88.4 ± 5.9. HR at peak exercise was 113.3 ± 17.3 beats/min (table 3).

The Outcome of ^{99m}SPECT

PBF was measured in the three parts of both lungs, and the middle lobes that had better PBF were estimated as the standard for each patient. The PBF scores were 1,105.9 ± 440.7 in the left upper lung, 1,338.9 ± 441.8 in the left middle lung, 1,098.6 ± 335.2 in the left lower lung, 1,187.8 ± 611.6 in the right upper lung, 1,498.1 ± 703.9 in the right middle lung, 1,570.8 ± 797.4 in the right lower lung. Percent changes of each lobe from the standard middle lung were 76.6 ± 28.1% in the left upper lobe, 89.6 ± 13.6% in the left middle lung, 80.3 ± 30.5% in the left lower lung, 94.7 ± 10.4% in the right middle lung and 102.2 ± 24.3% in the right lower lung. The mismatch scores of PBF were 1.1 ±,1.1 (table 4).

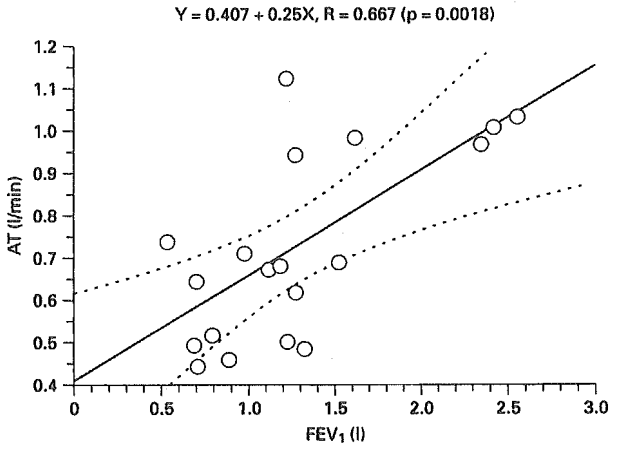


Fig. 5. FEV_1 correlated with AT (p = 0.0018). Patients with a higher FEV_1 showed a higher AT. The figure shows the 95% confidence bands for the true mean AT.

Correlation between the Indices of PFT, Exercise Test, CT Grade and ^{99m}SPECT

We determined that the following correlations existed:

- FEV_1 significantly correlated with $\dot{V}O_2$ peak (r = 0.648, p = 0.002) (fig. 3).
- FEV_1 significantly correlated with AT (r = 0.667, p = 0.0018) (fig. 5).

Table 3. Outcome of treadmill exercise and CT grading

Name	$\dot{V}O_2$ peak l/min	%$\dot{V}O_2$ peak %	%$\dot{V}CO_2$ peak	$\dot{V}_E$ peak l/min	$\dot{V}O_2$ AT l/min	$\dot{V}CO_2$ AT l/min	$\dot{V}_E$ AT	SaO_2 peak %	HR peak beats/min	Peak exercise time, min	CT grade
HA	1.223	105.7	1.34	53.48	0.988	0.984	35.73	89	137	4.73	1
SS	0.561	42.4	0.536	23.81	0.446	0.411	19.95	87	77	2.75	3
YS	0.92	61.7	0.819	33.58	0.742	0.582	26.47	90	126	4.1	2
IH	0.803	65.8	0.746	39.97	NA	NA	NA	96	90	3.3	1
TH	0.641	79.7	0.576	28.18	0.495	0.443	23.09	77	120	2.13	2
NH	0.677	32.1	0.71	42.91	0.486	0.517	36.6	96	77	1.6	2
MT	1.372	84.1	1.494	82.71	1.013	0.902	45.28	85	114	5.05	3
DY	0.858	44.4	0.886	36.14	0.69	0.688	29.54	78	124	3.55	4
AT	0.583	46.2	0.515	18.72	0.522	0.456	19.1	80	124	1.35	3
KK	1.315	77.5	1.195	40.18	0.683	0.578	23.02	94	120	3.45	3
IM	0.761	55.9	0.819	32.3	0.644	0.651	28.44	88	90	1.6	0.5
SI	1.301	151.5	1.223	43.85	1.123	1.003	39.57	94	130	2.5	0.5
YJ	0.867	55.5	0.788	33.49	0.621	0.545	24.94	82	116	2.2	3
IH	1.027	94.8	0.97	30.18	0.674	0.611	22.2	86	105	3.25	2
FK	1.108	78.2	1.096	30.62	0.946	0.903	28.02	84	111	1.9	0.5
TH	1.213	73.8	1.272	45.46	0.971	0.868	34.07	88	126	4.3	3
KO	0.638	85.5	0.609	29.4	0.464	0.447	23.03	93	125	2.23	1
YM	0.835	80.1	0.792	36.4	0.503	0.477	25.71	94	125	2.35	3
HS	0.764	82.0	0.66	24.7	0.713	0.577	22.23	92	107	2.85	1
MT	1.385	84.3	1.228	43.18	1.036	0.864	32.16	94	121	4.6	1
Mean	0.943	74.1	0.914	37.46	0.724	0.658	28.38	88.4	113.3	2.99	2.0
SD	0.278	26.4	0.295	13.57	0.222	0.199	7.156	5.9	17.3	1.12	1.1

NA = Not available.

Table 4. Outcome of ^{99m}SPECT

Name	Left upper	Left middle	Left lower	Right upper	Right middle	Right lower	Left upper %	Left middle %	Left lower %	Right upper %	Right middle %	Right lower %	PBF mismatch score
HA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SS	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
YS	1,404	1,998	951	1,861	1,940	1,989	70.3	100	47.6	93.1	97.1	99.5	1
IH	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
TH	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NH	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
MT	1,056	1,894	1,900	1,090	1,550	1,401	55.8	100	100.3	57.6	81.8	74	2
DY	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
AT	491	1,700	788	731	1,752	855	28	97	45	41.5	100	48.8	4
KK	1,076	1,166	891	1,025	1,236	1,289	87.1	94.3	72.1	82.9	100	104.3	0
IM	1,468	1,788	1,561	2,214	2,797	3,780	52.5	63.9	55.8	79.2	100	135.1	2
SI	1,710	1,800	1,213	1,879	2,471	2,063	69.2	72.8	49.1	76	100	83.5	1
YJ	506	1,073	1,519	865	1,254	1,802	40.4	85.6	121.1	69	100	143.7	2
IH	497	507	716	304	515	661	96.5	98.4	139	59	100	128.3	1
FK	1,294	1,205	1,086	988	758	1,095	107.4	100	90.1	82	62.9	90.9	0
TH	907	1,125	1,237	599	1,348	1,547	67.3	83.5	91.8	44.4	100	114.8	1
KO	1,615	1,149	833	1,542	1,115	1,070	140.6	100	72.5	134.2	97	93.1	1
YM	689	721	835	479	736	734	93.6	98	113.5	65.1	100	99.7	0
HS	1,808	1,600	920	2,189	2,623	2,471	68.9	61	35.1	83.5	100	94.2	1
MT	961	1,018	930	863	879	1,234	94.4	100	91.4	84.8	86.3	121.2	0
Mean	1,105.9	1,338.9	1,098.6	1,187.8	1,498.1	1,570.8	76.6	89.6	80.3	75.2	94.7	102.2	1.1
SD	440.7	441.8	335.2	611.6	703.9	797.4	28.1	13.6	30.5	22.1	10.4	24.3	1.1

NA = Not available.

Table 5. Correlation among the indices of grading of HRCT, ^{99m}SPECT and treadmill exercise tests

	r	p
CT vs. $\%\dot{V}O_2$ peak	−0.467	0.0689
CT vs. PBF	0.266	0.3376
$\%\dot{V}O_2$ peak vs. PBF	−0.327	0.2377
SaO_2 at peak exercise vs. PBF	−0.458	0.0985
CT vs. $FEV_1\%$	−0.064	0.7818
CT vs. $\%FEV_1$	−0.031	0.8931
CT vs. FEV_1	0.135	0.5552
CT vs. %VC	0.259	0.2596
CT vs. VC	0.323	0.1592

- CT grade did not correlate with PBF mismatch score ($r = 0.266$, $p = 0.3376$).
- $\%\dot{V}O_2$ peak did not correlate with CT grade ($r = -0.467$, $p = 0.0689$) or
- PBF mismatch score ($r = -0.327$, $p = 0.2377$).
- PBF mismatch score did not correlate with SaO_2 at peak exercise ($r = -0.458$, $p = 0.0985$).
- CT grade did not correlate with$FEV_1\%$ ($r = -0.064$, $p = 0.7818$), $\%FEV_1$ ($r = -0.031$, $p = 0.8931$), FEV_1 ($r = 0.135$, $p = 0.5552$), %VC ($r = 0.259$, $p = 0.2596$) and VC ($r = 0.323$, $p = 0.1592$) (table 5).
- CT grade did not correlate with TLC ($r = 0.192$, $p = 0.4275$), FRC ($r = 0.263$, $p = 0.2789$) or RV/TLC ($r = -0.126$, $p = 0.6037$).
- $\%DL_{CO}$ was measurable in 14 patients.
- $\%DL_{CO}$ did not correlate with $\dot{V}O_2$ peak ($r = 0.472$, $p = 0.088$).

Discussion

From our results, we learned that FEV_1 and VC correlated with $\dot{V}O_2$ peak and AT. This means that the indices of PFT predict exercise tolerance in emphysematous patients. This result is compatible with the report by Babb et al. [18].

They reported that $\%FEV_1$ correlated with $\dot{V}O_2$ max in patients with chronic airflow limitation [18]. The fact that CT grade did not correlate with TLC, FRC or RV/TLC might be due to the low number of measurable samples, and, because we adopted the visual methods of CT grading, we might have underestimated lung volume. However, there is a report that the morphological change of HRCT correlate with lung volume and DL_{CO} [19]. CT was reported to be useful not only for the evaluation of emphysematous change but also for the assessment of patients undergoing either lobe resection for a tumor or lung volume reduction surgery [20].

Our results indicate that emphysematous changes evaluated by HRCT do not correlate with the indices of $\%\dot{V}O_2$ peak or PBF mismatch score, nor does $\%\dot{V}O_2$ peak correlate with PBF mismatch score. Our results did not support statistically that the indices of HRCT or ^{99m}SPECT predicted exercise tolerance. A classical index of PFT, FEV_1, VC was well correlated with $\dot{V}O_2$ peak. The quantification of CT results in emphysematous patients was recently reported. For example, Coxon et al. [20] reported that CT-predicted surface to volume ratio correlated with histology, DL_{CO} [20]. These quantifications of CT analysis should be included in our analyses. LAA% is understood to be the most accurate index in quantifying emphysema [6]. However, the visual method and grading of CT that we adopted were reported to reflect the same change of macroscopic emphysema [12–14, 19].

In each case, it is useful for the assessment of emphysematous patients to follow up with CT and ^{99m}SPECT. Among the parameters we examined, it was difficult to find any correlation between the indices of CT grade, PBF mismatch score and exercise tolerance. However, HRCT was useful for detecting early changes of emphysema. Uppaluri et al. [21] reported that HRCT is useful in the diagnosis of regional diseases, such as pulmonary fibrosis and sarcoidosis. Commonly, pulmonary perfusion scans are useful for detecting pulmonary thromboembolism. Especially, ^{99m}SPECT increased the detection rate of lung cancer not only in emphysematous patients but also in cases of lung cancer [11]. Pulmonary blood flow decreased in the region of pulmonary emphysema. PBF mismatch score did not significantly correlate with SaO_2 at peak exercise (table 5). The increasingly inappropriate distribution of PBF may cause hypoxemia in some parts.

HRCT is a noninvasive method for the reliable detection of pulmonary emphysema. It is, however, associated with a considerable radiation dose for the patients. HRCT is not suitable for epidemiological or occupational studies [22], although, CT scanning is also used to diagnose the severity of bronchiectasis and its emphysematous changes [23]. At the moment, HRCT is the most precise instrument for the detection of anatomic emphysema, although it is difficult to discriminate anatomic emphysema from COPD. Therefore, PFT was helpful for the evaluation of PBF in emphysematous or COPD patients.

In this study, a significant correlation was perhaps not obtained because PFT and treadmill exercise tests are effort-dependent examinations, while HRCT and

^{99m}SPECT are static nonindependent examinations requiring no effort. If more data had been available, the correlation might have been different.

Clinically, HRCT is useful for detecting emphysematous changes, and the correlation between the CT grade and some indices of PFT was reported to be significant [12–14]. In addition, the indices of PFT were reported to significantly correlate with the indices of exercise tolerance. Therefore, the emphysematous changes detected by HRCT may predict the deterioration of exercise tolerance in persons with chronic pulmonary emphysema. Also, an increase in mismatched PBF reflects a decrease in the pulmonary vascular area, so the aggravation of mismatched PBF is useful for detecting the progression of emphysema, since the decrease in PBF reflects increasing emphysematous changes in emphysematous patients [10].

CT is a useful method for quantifying the extent of emphysema and detecting the progression in emphysematous changes [20]. Besides HRCT, there are now new methods of diagnosis: Kohlhäufl et al. [22] have recently reported aerosol-derived airway morphometry to diagnose chronic pulmonary emphysema morphologically. Even without CT and ^{99m}SPECT, the indices of PFT from a spirometer could be a useful marker of exercise performance in emphysematous patients. If an emphysematous patient has the additional complication of bronchiectasis, HRCT is an advantageous method of diagnosis [23], and if an emphysematous patients has bullae and needs bullectomy, CT is an advantageous method of diagnosis [24]. Conventional PFT may be the most useful tests available for the detection of aggravation in the exercise performance of emphysematous patients.

References

1 Gevenois PA, De Vuyst P, de Maertelaer V, Zanen J, Jacobovitz D, Cosio MG, Yernault JC: Comparison of computed density and microscopic morphometry in pulmonary emphysema. Am J Respir Crit Care Med 1996;154: 187–192.

2 Fujimoto S, Hirata K, Kurihara N: Exercise tolerance and limiting factor. Respir Circ 1996; 44:345–351.

3 Fujita E, Nagasaka Y, Hazu R Ikeshita K, Fukuoka M: Correlation of peak expiratory flow rate with exercise tolerance in stable asthmatic patients. Jpn J Appl Physiol 1998;28: 273–283.

4 Murayama M: Japanese standard value of respirocirculatory index on exercise. Jpn Circ J 1992;56(suppl):1514–1523.

5 Vlahovic G, Russell ML, Crapo JD, et al: Cellular and connective tissue changes in alveolar septal walls in emphysema. Am J Respir Crit Care Med 1999;160:2086–2092.

6 Hitsuda Y, Kawasaki Y, Igishi T, Yasuda K, Kato K, Matsumoto S, Nakamoto M, Hashimoto K, Fujii Y, Miyata M, Sasaki T, Shimizu E: A study of the minimum number of slices required for quantification of pulmonary emphysema by computed tomography. Jpn J Respir Soc 2000;38:430–436.

7 Fujita E, Nagasaka Y, Hazu R, Ikeshita K, Fukuoka M: Comparison of bronchial asthma and emphysema with respect to treadmill exercise test. 5th West-Pacific Allergy Symposium 7th Korea-Japan Joint Allergy Symposium, 1997, pp 207–209.

8 Biernacki W, Redpath AT, Best JJK, MacNee W: Measurement of CT lung density in patients with chronic asthma. Eur Respir J 1997;10: 2455–2459.

9 Troosters G, Gosselink R, Decramer M: Six minute walking distance in healthy elderly subjects. Eur Respir J 1999;14:270–274.

10 Jones NL, Killian KJ: Exercise limitations in health and disease. N Engl J Med 2000;343: 632–641.

11 Isawa T: Pulmonary Nuclear Medicine. Sendai, Kinkoudou, 1995, pp 64–136.

12 Yamagishi M, Mori M, Hirai H , Mori Y, Koba H, Suzuki A: Correlation of grading of pulmonary emphysema by computed tomography to pulmonary function. Jpn J Thorac Dis 1988; 26:109–115.

13 Foster WL, Pratt PC, Roggki VL, Godwin JD, Halvorsen RA, Putman CE: Centrilobular emphysema: CT-Pathologic correlation. Radiology 1986;159:27.

14 Bergin C, Muller NL, Nicholas DM, Lillington G, Hogg JC, Mulllen B, Grymaloski MR, Obsborne S, Pare PD: The diagnosis of emphysema: A computed tomographic pathologic correlation. Am Rev Respir Dis 1986;133:541.

15 Nagasaka Y, Fujita E, Hazu R, Fukuoka M, Ebara H, Ozuka K, Fujii K: Exercise capacity is determined by mismatch of structural changes and perfusion in chronic pulmonary emphysema. Am J Respir Crit Care Med 1999;159: A477.

16 Beaver WL, Wasserman K, Whipp BJ: A new method for detecting anaerobic threshold by gas exchange. J Appl Physiol 1986;60:2020–2027.

17 Jones NL (ed): Clinical use of exercise testing. Clinical Exercise Testing. Philadelphia, Saunders, 1988, pp 78–166.

18 Babb TG, Long KA, Rodarte JR: The relationship between maximal expiratory flow and increases of maximal exercise capacity with exercise training. Am J Respir Crit Care Med 1997; 156:116–121.

19 Gibson GJ, MacNee W: Chronic obstructive pulmonary disease: Investigations and assessment of severity; in Postma DS, Siafakas NM (eds): Management of Chronic Obstructive Pulmonary Disease. European Respiratory Monograph 7. Sheffield, European Respiratory Society Journals, 1998, vol 3, pp 25–35.

20 Coxon HO, Rogers RM, Whittall KP, D'Yachkova Y, Pare PD, Sciurba FC, Hogg JC: A quantification of the lung surface area in emphysema using computed tomography. Am J Respir Crit Care Med 1999;159:851–856.

21 Uppaluri R, Hoffman EA, Sonka M et al: Computer recognition of regional lung diseases patterns. Am J Respir Crit Care Med 1999;160: 648–654.

22 Kohlhäufl M, Brand P, Selzer T, Scheuch G, Meyer Th, Weber N, Schulz H, Häussinger K, Heyder J: Diagnosis of emphysema in patients with chronic bronchitis: A new approach. Eur Respir J 1998;1:793–798.

23 Roberts HR, Wells AU, Milne DG, Rubens MB, Kolbe J, Cole PJ, Hansell DM. Air flow obstruction in bronchiectasis. Correlation between computed tomography features and pulmonary function tests. Thorax 2000;55:198–204.

24 Celli BR, Gordon LS, Heffner J, Tiep B, Ziment I, Make B, Braman S, Olsen G, Phyllips Y: in Izumi T (ed): Standards for the Diagnosis and Care of Patients with Chronic Obstructive Pulmonary Disease (in Japanese). Tokyo, Life Science Publishing, 1996, pp 8–13.

British Journal of Cancer (2003) 88, 335–341

www.bjcancer.com

Randomised phase III trial of irinotecan combined with cisplatin for advanced non-small-cell lung cancer

S Negoro*[,1], N Masuda[2], Y Takada[3], T Sugiura[4], S Kudoh[5], N Katakami[6], Y Ariyoshi[7], Y Ohashi[8], H Niitani[9] and M Fukuoka[10] for the CPT-11 Lung Cancer Study Group West

[1]Department of Clinical Oncology, Osaka City General Hospital, 2-13-22, Miyakojima-hondori, Miyakojima-ku, Osaka 534-0021, Japan; [2]2nd Department of Internal Medicine, Osaka Prefectural Habikino Hospital, Osaka 583-8588, Japan; [3]Department of Respiratory Diseases and Radiology, Hyogo Medical Center for Adults, Hyogo 673-0021, Japan; [4]Department of Respiratory Diseases, Aichi Cancer Center, Aichi 464-0021, Japan; [5]1st Department of Internal Medicine, Osaka City University Medical School, Osaka 545-8585, Japan; [6]Department of Pulmonary Diseases, Kobe City General Hospital, Hyogo 650-0046, Japan; [7]1st Department of Respiratory Diseases, Aichi Prefectural Hospital, Aichi 444-0011, Japan; [8]School of Health Sciences and Nursing, Faculty of Medicine, University of Tokyo, Tokyo 113-0033, Japan; [9]The Tokyo Cooperative Oncology Group, Tokyo 105-0012, Japan; [10]Department of Medical Oncology, Kinki University School of Medicine, Osaka 589-8511, Japan

To determine a standard combination chemotherapy for patients with advanced non-small-cell lung cancer (NSCLC), we conducted a phase III trial of irinotecan (CPT-11) to test the hypotheses that CPT-11+cisplatin is superior to cisplatin+vindesine and that CPT-11 monotherapy is not inferior to cisplatin+vindesine. A total of 398 patients with previously untreated NSCLC were randomised to receive cisplatin+CPT-11 (CPT-P), cisplatin+vindesine (VDS-P) or CPT-11 alone (CPT). In the CPT-P arm, CPT-11 60 mg m^{-2} was administered on days 1, 8 and 15, and cisplatin 80 mg m^{-2} was administered on day 1. In the VDS-P arm, cisplatin 80 mg m^{-2} was administered on day 1, and vindesine 3 mg m^{-2} was administered on days 1, 8 and 15. In the CPT arm, CPT-11 100 mg m^{-2} was administered on days 1, 8 and 15. The median survival time was 50.0 weeks for patients on CPT-P, 45.6 weeks for those on VDS-P and 46.0 weeks for those on CPT ($P=0.115$, CPT-P *vs* VDS-P; $P=0.089$, CPT *vs* VDS-P), and the hazard ratio was 0.85 (95% confidence interval (CI): 0.65–1.11) for CPT-P *vs* VDS-P and 0.83 (0.64–1.09) for CPT *vs* VDS-P. The response rate was 43.7% for patients on CPT-P, 31.7% for those on VDS-P and 20.5% for those on CPT. Major adverse reactions were grade 4 neutropenia observed in 37, 54 and 8% of the patients on CPT-P, VDS-P and CPT, respectively; and grades 3 and 4 diarrhoea observed in 12, 3 and 15% of the patients, respectively. CPT-P therapy produces comparable survival to VDS-P in patients with advanced NSCLC. CPT-11 monotherapy is not inferior to VDS-P in terms of survival. The CPT-11-containing regimen is one of the most efficacious and well tolerated in the treatment of advanced NSCLC.
British Journal of Cancer (2003) **88**, 335–341. doi:10.1038/sj.bjc.6600725 www.bjcancer.com

Keywords: irinotecan; non-small-cell lung cancer; randomised phase III trial

In Japan, lung cancer is the leading cause of death among all the cancers, accounting for approximately 50 000 deaths annually (Statistics and Information Department, Minister's secretariat, Ministry of Health and Welfare, 2000). Non-small-cell lung cancer (NSCLC) accounts for more than 80% of primary lung cancers. Approximately two-thirds of NSCLC patients have advanced-stage cancer at presentation. The median survival time (MST) of the patients with advanced NSCLC, attained using the best available therapy (such as cisplatin-based chemotherapy), is typically 6–10 months, and most patients die of cancer within 1–2 years of diagnosis.

Irinotecan hydrochloride (CPT-11) is a water-soluble derivative of camptothecin, an alkaloid originally extracted from the Chinese tree *Camptotheca acuminata*. Differing from conventional anti-tumour drugs in mechanism of action, CPT-11 produces its effect by inhibiting the synthesis of DNA and RNA through inhibition of DNA topoisomerase I (Kawato *et al*, 1991). CPT-11 has shown strong antitumour activity as a single agent against a broad spectrum of experimental tumours as well as against human malignancies.

The phase I clinical trial of CPT-11, in which CPT-11 was administered weekly, showed that the dose-limiting adverse effects were leucopenia and diarrhoea, and the recommended dose for the phase II monotherapy trial was 100 mg m^{-2} week^{-1} (Negoro *et al*, 1991). In the phase II clinical trial of CPT-11 monotherapy for patients with untreated advanced NSCLC, the response rate was 31.9% (23 out of 72 cases) and the MST was 42 weeks (Fukuoka *et al*, 1992).

In preclinical studies, CPT-11 was confirmed to act synergistically with cisplatin (CDDP) (Kudoh *et al*, 1993). Since CDDP is active against NSCLC and forms part of the standard therapeutic armamentarium employed in the treatment of NSCLC, a phase I trial of CPT-11+CDDP was initiated in 1991 (Masuda *et al*, 1992). From this trial, '60 mg m^{-2} CPT-11 on days 1, 8 and 15, and 80 mg m^{-2} CDDP on day 1' were recommended as the optimal dose

*Correspondence: Dr S Negoro;
E-mail: m6122765@msic.med.osaka-cu.ac.jp
Received 3 April 2002; revised 18 October 2002; accepted 28 October 2002

schedule. In 1992, a phase II trial of CPT-11+CDDP for untreated, stage IIIB or IV NSCLC was started employing this dose schedule; the response rate was 52% and the MST was 44 weeks (Masuda *et al*, 1998). These results suggested that CPT-11+CDDP could improve outcomes for patients with advanced NSCLC.

Previously, the National Cancer Institute of Canada (NCIC) Clinical Trials Group undertook a randomised study involving patients with NSCLC to determine whether there was a survival advantage for those patients treated with combination chemotherapy (vindesine (VDS)+CDDP or cyclophosphamide+doxorubicin+CDDP) over best supportive care (BSC). The results of that trial showed that the patients receiving VDS+CDDP had a statistically significant survival advantage over patients receiving BSC (Rapp *et al*, 1988). The VDS+CDDP regimen is one of the most widely used chemotherapeutic regimens for advanced NSCLC in Japan, and it was considered to be the most appropriate control regimen at the time when this study was initiated.

Based on the above results, we planned a phase III trial to compare CPT-11+CDDP and CPT-11 alone, with the control arm of VDS+CDDP, in order to elucidate the role of CPT-11 in advanced NSCLC.

MATERIALS AND METHODS

Eligibility criteria

Patients with histologically or cytologically confirmed and previously untreated NSCLC were enrolled into this trial. Patients with stage IIIB or IV cancer were eligible if they had measurable disease and a life expectancy of at least 3 months. Additional inclusion criteria were age (15–75 years old); a performance status (PS) of 0–2 on the Eastern Cooperative Oncology Group (ECOG) scale; and adequate functional indices for bone marrow (leucocyte count between $\geqslant$4000 and $<$12 000/mm^{-3}, platelet count $\geqslant$100 000/mm^{-3}, haemoglobin concentration $\geqslant$9.5 g dl^{-1}), liver (GOT and GPT $<$100 IU l^{-1}, serum bilirubin $\leqslant$1.5 mg dl^{-1}), kidneys (serum creatinine $\leqslant$ the upper limit of normal) and lungs (PaO_2; at rest $\geqslant$8.0 kPa). Patients with other concurrent malignancies or a history of other malignancies, active infection, diarrhoea (watery stool), paralytic ileus, pulmonary fibrosis, pericardial effusion, considerable pleural effusion or ascites, uncontrolled diabetes mellitus or symptomatic metastasis to the brain, were excluded. Informed consent was obtained from each patient before enrollment. Each institutional review board for human experimentation approved the protocol of this study.

Randomisation

Eligible patients were randomised to one of the three treatment arms by a centralised dynamic balancing method (a modified minimisation method) using stage (IIIB/IV), PS (0–1/2) and institution as balancing variables (Figure 1).

Treatment schedule

In the CPT-P arm, CPT-11 was given intravenously (i.v.) on days 1, 8 and 15 at a dose of 60 mg m^{-2}, and CDDP was given i.v. on day 1 at a dose of 80 mg m^{-2}. In the VDS-P arm, CDDP was given i.v. on day 1 at a dose of 80 mg m^{-2}, and VDS was given i.v. on days 1, 8 and 15 at a dose of 3 mg m^{-2}. In the CPT arm, CPT-11 was given i.v. on days 1, 8 and 15 at a dose of 100 mg m^{-2}. In each arm, one course of treatment lasted 4 weeks and each course was repeated more than twice, until occurrence of unacceptable toxicity, disease progression, patient's refusal and investigator's medical decision.

CPT-11, diluted in $\geqslant$500 ml of normal saline, was administered by i.v. infusion over 90 min. VDS was administered as an i.v. push with a running of 5–10 ml of normal saline. CDDP was administered i.v. as an undiluted solution and infused for a period of 60 min with more than 2600 ml of hydration and a diuretic before and after administration. To control CDDP-induced emesis, a 5-HT_3 receptor antagonist was given before CDDP administration. In the CPT-P arm, CDDP was administered after the CPT-11 infusion.

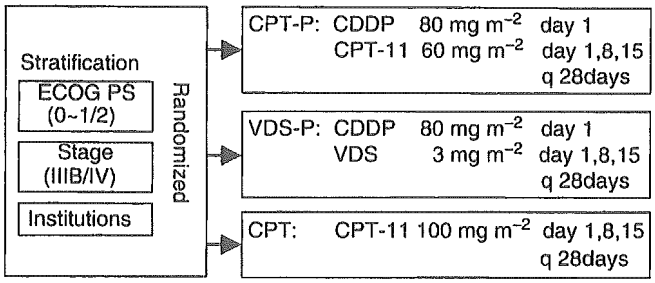


Figure 1 Study design. Patients enrolled were randomly allocated to receive CPT-P, VDS-P or CPT, after being stratified by PS, disease stage and institution. CPT-11 = irinotecan; CDDP = cisplatin; VDS = vindesine.

Dose modification

In all three arms, the dose of CPT-11 or VDS on day 8 or 15 was not given if the leucocyte count was less than 3000 mm^{-3}, platelet count was less than 100 000 mm^{-3}, body temperature was 37.5°C or higher, or diarrhoea was grade 1 (frequency increased to two times or more daily, abdominal pain rated as mild or severe, or watery stool) or higher on the ECOG scale.

In each arm, before the next course was started, the leucocyte count had to be 4000 mm^{-3} or more, platelet count had to be 100 000 mm^{-3} or more, serum creatinine concentration had to be normal, and diarrhoea and fever should have disappeared. If there was a delay greater than 4 weeks caused by persistent toxicity, patients in each arm were to be withdrawn from the study.

If during the previous course, the leucocyte count had been less than 1000 mm^{-3}, platelet count less than 50 000 mm^{-3} or diarrhoea had been grade 2 (stool passage increased to four times or more daily, stool passage at night or moderate abdominal pain, watery stool) or higher on the ECOG scale, the dose of CPT-11 was reduced to 50 mg m^{-2} in the CPT-P arm, and to 80 mg m^{-2} in the CPT arm. If during the previous course the serum creatinine concentration was more than 1.5 times the upper normal limit, the dose of CDDP was reduced to 60 mg m^{-2} in the CPT-P arm and in the VDS-P arm.

Evaluation

This trial was independently monitored and performed according to GCP rules.

The primary end point of this study was overall survival. Response rate, duration of response, time to disease progression and toxicity were all secondary end points.

The stage of the disease was determined based on a complete medical history and physical examination, routine chest radiographs, fibreoptic bronchoscopy, computed tomography of the head, chest and abdomen and bone scintigraphy. The staging estimates were made according to the international staging system (Mountain, 1986). Before the first course, the haemogram of each patient was determined and serum chemistry was used to check renal and hepatic functions, electrolytes and urinalysis. The haemogram was assessed at least twice a week, and serum chemistry, electrolytes, urinalysis and chest radiographs at least once weekly. Computed tomography of the chest was repeated monthly. Any other examination was carried out when any clinical sign of disease progression was observed. The eligibility, evaluability and response of each patient were reviewed extramurally.

The assessment of antitumour effects and toxicity was based on the WHO criteria (World Health Organization, 1979), but diarrhoea was evaluated in accordance with the ECOG common toxicity criteria (Oken *et al*, 1982). The overall survival was defined as the time from randomisation to the time of death from any cause or to last follow-up (cutoff date: 23 January 2000). The duration of each response was defined as the time from the initial response to disease progression. The time to disease progression was defined as the time from the start of treatment to disease progression.

Health-related quality of life (QOL) was assessed experimentally with a self-administered QOL questionnaire for cancer patients treated with anticancer drugs (QOL-ACD) (Kurihara *et al*, 1999). The QOL-ACD was completed at baseline, every week during treatment and every month after treatment.

Statistical analysis

The superiority of CPT-P to VDS-P and the noninferiority of CPT to VDS-P were evaluated in terms of survival benefit.

The sample size was calculated on the basis that MSTs were expected to be 44, 30 and 35 weeks for the CPT-P, VDS-P and CPT arms, respectively (Einhorn *et al*, 1986), and both the accrual and follow-up intervals were 2 years.

By using the Schoenfeld and Richter equation (1982), the least number of patients to provide the 80% power needed to confirm the superiority of a regimen was calculated to be 115 per treatment arm for a one-sided 2.5% significance level test. Furthermore, the least number of deaths to provide the 80% power needed to prove the noninferiority for the upper limit at a 95% CI for the hazard ratio of CPT compared with VDS-P was lower than the upper equivalence margin, 1.33, and was estimated to be 81 per treatment arm at the one-sided 2.5% level. The significance levels for both inferences were set at 2.5% to control the overall type I error rate, and the one-sided statistical approach was employed to keep the simplicity and structural consistency of the statistical inferences for two study hypotheses. Taking ineligible patients into account, the sample size was set at 130 per treatment arm.

Cumulative survival curves were constructed as time-to-event plots by the Kaplan–Meier method (1958). Differences between the curves were tested for significance using one-sided log-rank statistics, and were estimated for noninferiority using the hazard ratio produced by the Cox regression model (Cox and Oakes, 1984). Furthermore, Cox regression models were used to evaluate treatment effects on survival, with adjustment for well-known prognostic factors: stage, PS, gender, weight loss, albumin and LDH (Albain *et al*, 1991; Espinosa *et al*, 1995; Paesmans *et al*, 1995; Ray *et al*, 1998). Response rates and toxicities were compared using the χ^2 test. Subgroup analysis for survival was conducted by stage.

RESULTS

Patient characteristics

From July 1995 to January 1998, 398 patients from 41 centres were entered into this study. Of the 398 patients randomised, 18 patients did not meet the eligibility criteria. The reasons for exclusion were: early stage (<IIIB) in five patients, previous treatment (OK-432, radiation therapy or surgery for local metastasis) in eight patients, other concurrent malignancies in two patients, age (>75 years) in one patient, presence of pericardial effusion in one patient, and leucocyte count ($>12\,000\,mm^{-3}$) in one patient.

The patient characteristics at baseline are listed in Table 1. The median age of the patients was 62–64 years and the range was 35–75 years. A total of 62% of the patients had adenocarcinoma, 31% had squamous cell carcinoma, 37% had stage IIIB disease and 94% of the patients had a good performance status (PS 0–1). The three treatment groups were well balanced for all baseline characteristics.

Table 1 Patient characteristics at baseline

	CPT-P	VDS-P	CPT
Patients entered	133	133	132
Eligible patients[a]	129	122	129
Gender			
Male	98	98	96
Female	31	24	33
Age (years)			
–49	10	19	22
50–59	36	29	30
60–69	58	51	52
70–	25	23	25
Median (range)	64 (36–75)	64 (35–75)	62 (35–75)
PS			
0–1	121	115	121
2	8	7	8
Stage			
IIIB	49	46	44
IV	80	76	85
Histology			
Adenocarcinoma	78	73	84
Squamous cell carcinoma	37	45	34
Others	14	4	11
Weight loss			
No	74	67	77
Less than 5%	10	14	9
5% or more	27	30	29
Unknown	18	11	14

[a]Patients who were used to evaluate survival.
CPT=irinotecan; P=cisplatin; VDS=vindesine; PS=performance status.

Treatment administration

Three patients received no treatment (one patient in the CPT-P arm and two patients in the CPT arm), two patients in the VDS-P arm received vincristine (VCR) instead of VDS, one patient in the CPT-P arm exceeded the daily dose of CPT-11 and CDDP, and one patient in the CPT-P arm received CDDP on days 1 and 8. These seven patients were included in the survival analysis, but excluded from response and toxicity analysis. In all, 109 of the 126 patients on CPT-P (87%), 97 of the 120 patients on VDS-P (81%) and 101 of the 127 patients on CPT (80%) received more than two courses of treatment, and 22 patients (18%) on CPT-P, 15 patients (13%) on VDS-P and 20 patients (16%) on CPT received more than four courses of treatment. The median number of courses administered per patient was 3 for CPT-P and 2 for both VDS-P and CPT. The number of treatment courses varied from one to six in each arm. In the CPT-P, VDS-P and CPT arms, 23, 10 and 16 patients were withdrawn from the study because of toxicities (including duplications); 28, 44 and 51 patients because of disease progression; 37, 27 and 19 patients due to patient's refusal; 10, 19, and 22 patients because of aggravated clinical symptom; 50, 32 and 28 patients due to investigator's medical decision; and 6, 9 and 9 patients because of other reasons, respectively. Most patients could repeat treatment courses every 28 days irrespective of the arm.

The number of treatment courses in which CPT-11 or VDS was administered on days 1, 8 and 15 totaled 164 courses (50% of a

total of 330 courses) for the CPT-P arm, 66 courses (24% of a total of 279 courses) for the VDS-P arm and 182 courses (62% of a total of 292 courses) for the CPT arm. Dose omissions on day 8 and/or day 15 were because of diarrhoea and/or leucopenia in the CPT-P arm and the CPT arm, and because of leucopenia in the VDS-P arm. The median dose intensity of CDDP was the same ($20\,mg\,m^{-2}\,week^{-1}$) for both the CPT-P and VDS-P arms. The median dose intensity of CPT-11 was $30\,mg\,m^{-2}\,week^{-1}$ (67% of planned dose) for the CPT-P arm and $61.3\,mg\,m^{-2}\,week^{-1}$ (82% of planned dose) for the CPT arm. The median dose intensity of VDS was $1.5\,mg\,m^{-2}\,week^{-1}$ (67% of planned dose).

Survival

The survival curves for all eligible patients are shown in Figure 2. The MST was 50.0 weeks for patients on CPT-P, 45.6 weeks for those on VDS-P and 46.0 weeks for those on CPT, and the 1- and 2-year survival rates were 46.5 and 19.4% for patients on CPT-P, 38.3 and 18.7% for those on VDS-P, and 41.8 and 21.9% for those on CPT. The one-sided log-rank test comparing the survival of patients who received CPT-P *vs* those treated with VDS-P yielded a *P*-value of 0.115, and when the survival of patients treated with CPT was compared with that of patients treated with VDS-P, the *P*-value was 0.089. The hazard ratio was 0.85 (95% CI: 0.65–1.11) for CPT-P *vs* VDS-P and 0.83 (95% CI: 0.64–1.09) for CPT *vs* VDS-P.

Subgroup analyses for survival were conducted using stage as one of the balancing variable factors (Figure 3). Among patients with stage IIIB disease (Figure 3A), the MST was 49.7 weeks for those on CPT-P; 60.6 weeks for those on VDS-P and 63.3 weeks for those on CPT. The hazard ratio was 1.24 (95% CI: 0.81–1.91, one-sided, log-rank test: $P=0.838$) for CPT-P *vs* VDS-P and 0.99 (95% CI: 0.63–1.56, one-sided, log-rank test: $P=0.483$) for CPT *vs* VDS-P.

Among patients with stage IV disease (Figure 3B), the MST was 50.0 weeks for those on CPT-P, 36.4 weeks for those on VDS-P and 42.1 weeks for those on CPT. The hazard ratio was 0.64 (95% CI: 0.46–0.89, one-sided, log-rank test: $P=0.004$) for CPT-P *vs* VDS-P and 0.70 (95% CI: 0.50–0.98, one-sided, log-rank test: $P=0.018$) for CPT *vs* VDS-P.

The Cox proportional hazards model was used to adjust for and determine the impact of prognostic factors and treatment on survival (Table 2). Predictive factors for improved survival included early clinical stage, no weight loss, normal LDH, normal albumin and better PS, each of which was significantly associated with longer survival.

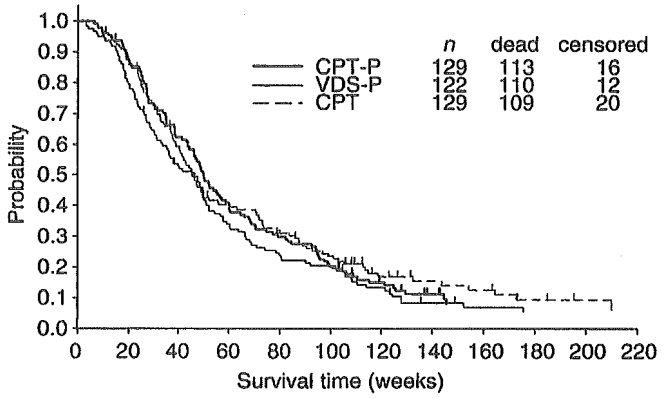


Figure 2 Survival of eligible patients. Survival time was calculated from the date patients were entered into this study. CPT = irinotecan; P = cisplatin; VDS = vindesine; *n* = number of eligible patients.

Response

Objective response data are listed in Table 3. The overall response rate was 43.7% for patients on CPT-P, 31.7% for those on VDS-P and 20.5% for those on CPT (two-sided, χ^2 test: $P<0.001$).

The median duration of response for responders was 141 days for patients on CPT-P, 121 days for those on VDS-P, and 117 days for those on CPT (two-sided, log-rank test: $P=0.601$). The median time to progression for the patients included in the efficacy analysis was 148 days for patients on CPT-P, 117 days for those on VDS-P and 100 days for those on CPT (two-sided, log-rank test: $P=0.091$).

Toxicity

Major adverse reactions are listed in Table 4. Grade 4 neutropenia was observed significantly more frequently in the VDS-P arm than in the CPT-P and CPT arms, both of which contain CPT-11 ($P < 0.001$). Grade 4 thrombocytopenia was more frequent in the CPT-P arm, but there were no thrombocytopenia-related complications.

Diarrhoea was the main nonhaematological sign of toxicity of the regimens containing CPT-11 (CPT-P and CPT); these two regimens were associated with a significantly higher occurrence of grades 3 and 4 diarrhoea as compared with the VDS-P regimen ($P=0.008$). The regimens containing CDDP (CPT-P, VDS-P) were associated with a significantly higher occurrence of grade 3 nausea/vomiting as compared with the CPT alone regimen ($P=0.001$). Peripheral neurological symptoms were observed significantly more frequently in the VDS-P group than in the other two groups ($P < 0.001$).

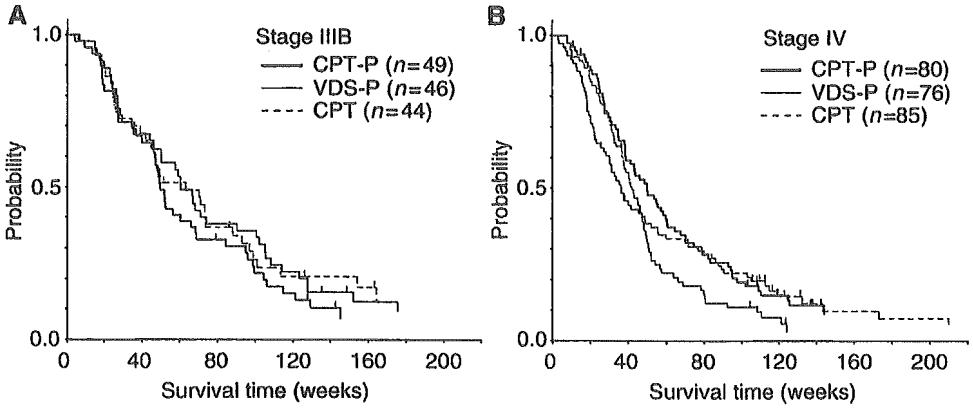


Figure 3 Survival of eligible patients: (**A**) stage IIIB, (**B**) stage IV. Survival time was calculated from the date patients were entered into this study. CPT = irinotecan; P = cisplatin; VDS = vindesine; *n* = number of eligible patients.