Original paper

# Cluster analysis of deterioration sites after first-line epidermal growth factor receptor-tyrosine kinase inhibitor in epidermal growth factor receptor mutated non-small cell lung cancer

**Shinichiro Okauchi[1], Kunihiko Miyazaki[2], Hiroaki Satoh[1]**

[1]Mito Medical Center, University of Tsukuba, Tsukuba, Japan
[2]Ryugasaki Saiseikai Hospital, Ryugasaki, Japan

**Introduction**: For epidermal growth factor receptor (EGFR) mutated non-small cell lung cancer (NSCLC), no studies have treated the site of recurrence after first-line tyrosine kinase inhibitor (TKI) treatment as a "metastasis pattern". This study aims to assess whether these patients have a specific "metastasis pattern" at the site of recurrence after the treatment.
**Material and methods**: Data were collected from all consecutive EGFR mutated NSCLC patients between 2009 and 2021. Metastatic patterns were analyzed using cluster analysis in patients with EGFR mutated NSCLC.
**Results**: During the study period, 83 EGFR mutated NSCLC patients were treated with EGFR-TKI. Patients who had no metastases at the time of diagnosis were divided into two groups according to the presence or absence of recurrence of metastases after TKI administration. Patients with metastases at diagnosis were divided into 4 groups by cluster analysis. A statistically significant difference in metastasis frequency was confirmed among these 6 groups ($\chi^2$ test, $p = 0.0001$). Furthermore, when the frequency of metastasis recurrence after TKI administration in these 6 groups was examined, a statistically significant difference was confirmed ($\chi^2$ test, $p = 0.0001$).
**Conclusions**: Even in EGFR mutation-positive patients, the knowledge of the recurrent patterns might be useful for clinical practice in the foreseeable future, as it enables more efficient detection of metastatic disease through imaging, and more effective treatment at predicted metastatic sites.

**Key words**: cluster analysis, deterioration, metastasis, recurrence, first-line EGFR-TKI, EGFR mutated NSCLC.

**Contemp Oncol (Pozn) 2022; 26 (2): 123–127**
**DOI: https://doi.org/10.5114/wo.2022.118195**

## Introduction

Cluster analysis is a multivariate analytic method that classifies target groups by creating clusters that are similar to each other, from groups of subjects with different properties [1, 2]. In clinical research, this type of analysis has been used in infectious diseases to type pathogens [3] and to explore the relationship between genotypes and phenotypes of bronchial asthma [4]. In lung cancer, it is hypothesized that there are 'specific metastatic patterns' rather than 'random' metastasis. According to this hypothesis, we conducted a study using cluster analysis, and showed the possibility of specific metastatic patterns in patients with epidermal growth factor receptor (EGFR) mutated non-small cell lung cancer (NSCLC) [5]. We also revealed that there were differences in metastatic patterns among patients with EGFR gene mutation NSCLC, SCLC, and squamous cell lung cancer [5].

Non-small cell lung cancer patients with EGFR mutations are expected to have a high response and long duration of response with EGFR-tyrosine kinase inhibitors (TKIs) [6]. However, it is considered difficult to cure by EGFR-TKI, and it will recur sooner or later after TKI treatment [6]. Metastases have been studied based on autopsy data. However, what is needed is information that is useful in clinical practice, and the usefulness of this information in the clinical setting is limited. Although most of the conventional analysis of metastasis has been to examine the frequency of metastasis in each organ, the frequency of each metastasis site is not sufficient information as clinical information suitable for each patient. It is not pointless to know the frequency of individual metastatic organs. However, if it is possible to investigate the pattern of metastatic site after first-line treatment, it may provide useful information for performing efficient imaging studies and selecting the appropriate treatment method. This time, we analyzed the recurrence pattern of distant organ metastasis after first-line EGFR-TKI treatment using cluster analysis in patients with EGFR mutated NSCLC, who are a group of patients expected to survive for a long time.

## Material and methods

### Patients

Patients who presented with pathologically diagnosed lung cancer between April 2009 and December 2021 at two tertiary hospitals in Japan – the University of Tsukuba Mito Medical Center and Ryugasaki Saiseikai General

*This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0). License (http://creativecommons.org/licenses/by-nc-sa/4.0/)*

Hospital – were identified retrospectively via computerized searches of tumor registry data. Medical record information from diagnostic imaging, including chest computed tomography (CT), brain magnetic resonance imaging or enhanced head CT, bone scan and ultrasonography and/or CT of the abdomen, was used to identify the location of metastatic tumors. Information on distant metastases was collected in detail, with the most common metastatic sites being lung, bone, brain, liver, adrenal gland, distant lymph nodes, and other sites. Clinical data for age, gender, smoking habit, primary site of lung cancer, maximum diameter of the primary tumor, and N-stage of lung cancer were also collected.

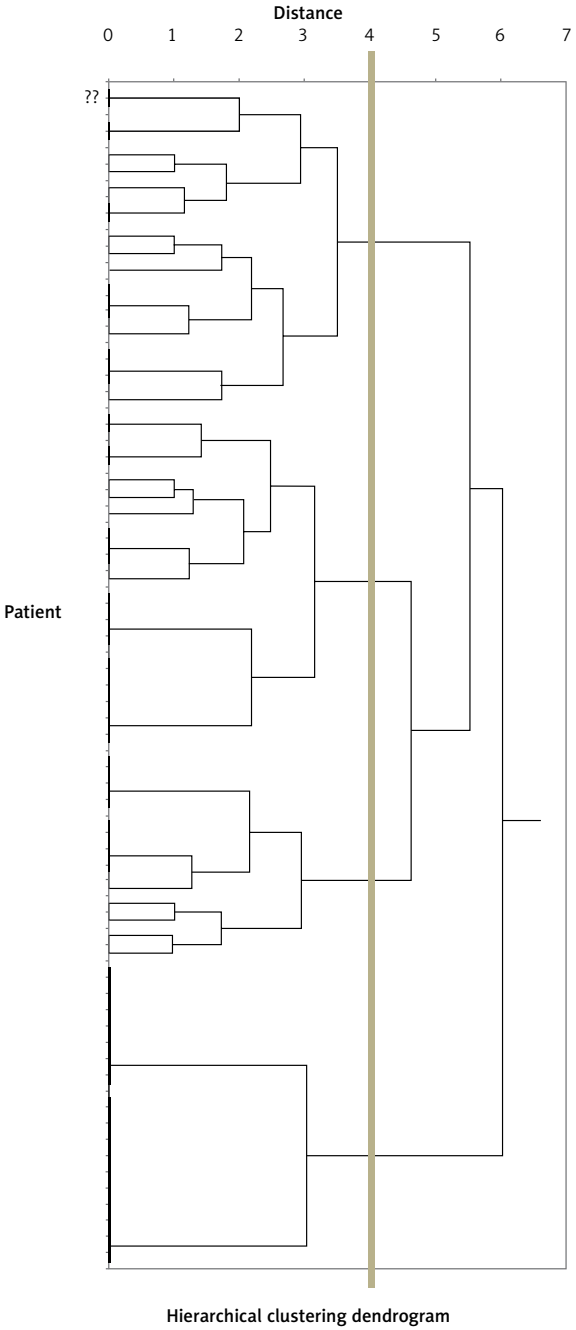


**Fig. 1.** Patients with stage IV A–B epidermal growth factor receptor mutated adenocarcinoma were divided into 4 clusters

## Ethics

This study was approved by the institutional ethics committee of each hospital. Written comprehensive consent was obtained from each patient regarding the use of the obtained clinical information for academic research including presentations at academic societies and publishing academic articles.

## Statistical analysis

Cluster analysis was performed to classify patients [1]. Briefly, pre-clusters to reduce the size of the matrix that contained the distances between all possible pairs of cases were performed. Then, the standard hierarchical clustering algorithm was applied to the pre-clusters to explore a range of solutions with different numbers of clusters. At this point, hierarchical cluster analysis was performed using Ward's method to generate a dendrogram for estimation of the number of likely clusters within the population. Cluster boundaries were defined by large differences between successive fusion levels [2]. At each cluster, samples were merged into larger clusters to minimize the within-cluster sum of squares or to maximize the between-cluster sum of squares in Euclidean distance. Variables for cluster analysis included the common metastatic sites described above. The type of EGFR mutation was also included as a variable in patients with EGFR mutant tumors. Statistical analyses were performed using BellCurve for Excel (version 3.0). Differences in proportions between two and among three independent groups were compared using the $\chi^2$ test. $P < 0.05$ was considered statistically significant.

## Results

There were 86 pathologically diagnosed NSCLC patients with EGFR mutation. Among them, 73 patients had distant metastasis and 13 patients had locally advanced NSCLC. The most common metastatic sites were the lung, pleura, bone, brain, liver, adrenal gland, lymph nodes other than regional nodes, and other sites. Figure 1 shows the dendrogram of 4 clusters created based on these metastatic sites in these 73 metastatic patients. In this cluster model, metastatic groups were identified as follows:

- cluster I (bone-other site type, number of patients – 20),
- cluster II (lung-other site type, number of patients – 21),
- cluster III (brain-other site type, number of patients – 13),
- cluster IV (pleura-lung type, number of patients – 19).

Demographic and baseline clinical and pathological characteristics of the identified clusters are shown in Table 1. There was a significant difference in frequency of distant organ metastasis in these 4 clusters of patients ($\chi^2$ test, $p = 0.0001$). Thirteen patients who had locally advanced NSCLC were divided into 4 patients who did not receive surgery and 9 patients who received surgery, and were designated as clusters V and VI, respectively.

We investigated the sites of recurrence after first-line EGFR-TKI treatment in these 6 cluster patients. Table 2 A shows the results of investigating new metastatic sites after the TKI treatment. There was a significant difference in frequency of newly developed distant or-

**Table 1.** Characteristics of patients in six clusters

| Variable | Cluster | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Number of patients | 20 | 21 | 13 | 19 | 4 | 9 |
| Age, median (range) | 73 (50–83) | 67 (54–83) | 69 (44–-86) | 75 (56–92) | 86 (79–88) | 77 (58–85) |
| Gender | 9/11 | 7/14 | 6/7 | 6/13 | 2/2 | 3/6 |
| EGFR, Ex19 del/Ex21 L858R/others | 9/8/3 | 10/11 | 9/4/0 | 9/8/2 | 2/1/1 | 6/2/1 |
| Stage, IIIB/IVA/IVB | 0/1/19 | 0/5/16 | 0/1/12 | 0/12/7 | 4/0/0 | 9/0/0 |
| Surgical resection, absent/present | 20/0 | 21/0 | 13/0 | 19/0 | 4/0 | 0/9 |
| Metastatic sites at diagnosis | | | | | | |
| Lung/pleura/brain | 1/5/1 | 21/2/10 | 0/1/13 | 8/19/0 | 0/0/0 | 0/0/0 |
| Bone/liver/adrenal gland | 17/6/8 | 9/3/0 | 4/4/2 | 0/0/0 | 0/0/0 | 0/0/0 |
| Lymph nodes/others | 5/5 | 1/0 | 1/1 | 0/0 | 0 0 | 0/0 |
| EGFR-TKI, O/A/others | 3/5/12 | 3/5/13 | 2/3/8 | 7/4/8 | 0/1/3 | 2/2/5 |

*A– afatinib, EGFR – epidermal growth factor receptor, Ex – exon, O – osimertinib, TKI – tyrosine kinase inhibitor*

**Table 2.** Newly emerged metastatic sites (**A**) and exacerbation site of metastasis at diagnosis and exacerbation site (**B**) after first-line tyrosine kinase inhibitor therapy

| A: Newly metastatic site | Cluster | | | | | | *p*-value* |
|---|---|---|---|---|---|---|---|
| | 1 (%) | 2 (%) | 3 (%) | 4 (%) | 5 (%) | 6 (%) | |
| Lung | 2 (9.5) | 3 (15.0) | 1 (7.7) | 6 (31.6) | 0 | 2 (22.2) | 0.0001 |
| Pleura | 6 (28.6) | 7 (35.0) | 6 (46.2) | 2 (10.5) | 2 (50.0) | 1 (11.1) | |
| Brain | 4 (19.0) | 1 (5.0) | 1 (7.7) | 4 (21.1) | 1 (25.0) | 2 (22.2) | |
| Bone | 9 (42.9) | 2 (10.0) | 6 (46.2) | 6 (31.6) | 1 (25.0) | 4 (44.4) | |
| Liver | 1 (4.8) | 0 | 2 (15.4) | 0 | 0 | 0 | |
| Adrenal gland | 0 | 0 | 0 | 0 | 0 | 1 (11.1) | |
| Lymph nodes | 1 (4.8) | 0 | 0 | 0 | 0 | 1 (11.1) | |
| Others | 5 (23.8) | 8 (40.0) | 3 (23.1) | 4 (21.1) | 0 | 1 (11.1) | |

| B: Exacerbation site | Cluster | | | | | | *p*-value* |
|---|---|---|---|---|---|---|---|
| | 1 (%) | 2 (%) | 3 (%) | 4 (%) | 5 (%) | 6 (%) | |
| Lung | 2 (9.5) | 5 (25.0) | 1 (7.7) | 7 (36.8) | 0 | 2 (22.2) | 0.0001 |
| Pleura | 9 (42.9) | 7 (35.0) | 7 (53.8) | 14 (73.7) | 2 (50.0) | 1 (11.1) | |
| Brain | 4 (19.0) | 4 (20.0) | 3 (21.3) | 4 (21.1) | 1 (25.0) | 2 (22.2) | |
| Bone | 14 (66.7) | 4 (20.0) | 7 (53.8) | 6(31.6) | 1 (25.0) | 4 (44.4) | |
| Liver | 2 (9.5) | 1 (5.0) | 2 (15.4) | 0 | 0 | 0 | |
| Adrenal gland | 0 | 0 | 0 | 0 | 0 | 1 (11.1) | |
| Lymph nodes | 1 (4.8) | 1 (5.0) | 0 | 0 | 0 | 1 (11.1) | |
| Others | 5 (23.8) | 8 (40.0) | 3 (23.1) | 4 (21.1) | 0 | 1 (11.1) | |

**Difference in ratio among six clusters ($\chi^2$ test)*

gan metastasis in these 6 clusters of patients ($\chi^2$ test, $p$ = 0.0001). Table 2 B shows the results of examining the exacerbation sites, focusing not only on the newly developed metastatic sites but also on the exacerbation of the lesions diagnosed at the time of initial presentation. Figure 2 shows the frequency of metastatic sites. There was a significant difference in frequency of the exacerbation sites in these 6 clusters of patients ($\chi^2$ test, $p$ = 0.0001).

## Discussion

Cluster analysis is one of the most useful statistical analytical methods, but it is a relatively new method of analysis in clinical oncology [7, 8]. Recently we studied metastatic patterns using cluster analysis at the time of initial diagnosis of NSCLC patients with EGFR mutation, small cell lung cancer, and squamous cell lung cancer [5, 9]. In those studies, metastatic groups were identified as follows: lung-liver type, lung-brain type, and bone-liv-

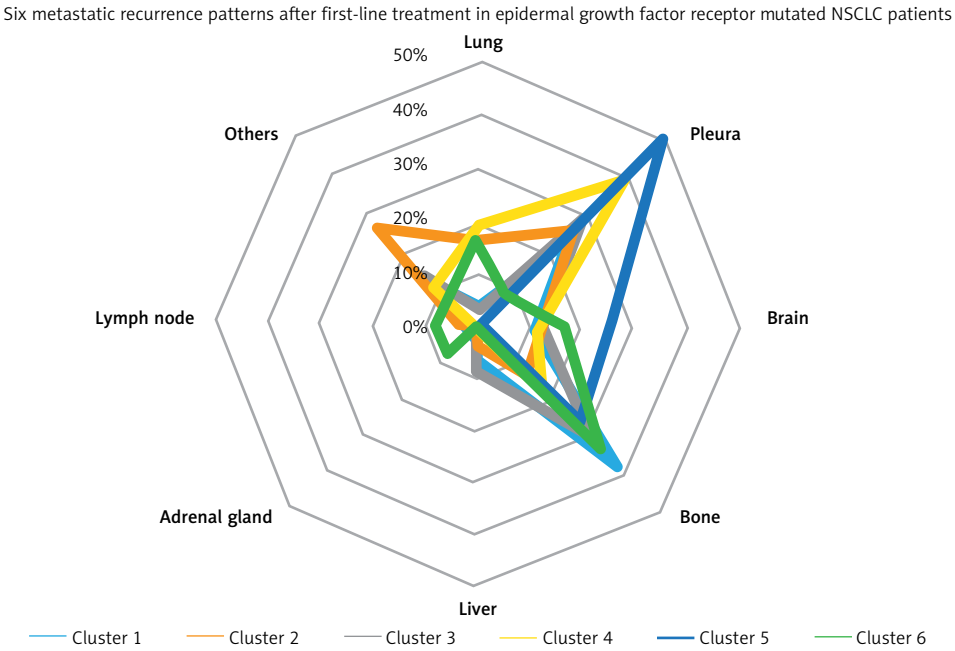


**Fig. 2.** The newly developed/deteriorated metastatic patterns after first-line epidermal growth factor receptor (EGFR)-tyrosine kinase inhibitor therapy in patients with EGFR mutated adenocarcinoma

er-other sites type [5]. We also found that patients with the bone-liver-other sites type had a poor prognosis [9]. In patients with EGFR gene mutations, EGFR-TKIs are usually selected as the first-line standard therapeutic agent [10]. However, the mainstream idea is that EGFR mutated NSCLC cannot be completely cured with currently available TKIs [11]. If so, recurrence is inevitable sooner or later, and if it is possible to infer the pattern of recurrence, efficient imaging testing might be possible. It was speculated that useful information could be obtained if the metastasis pattern could be inferred from the viewpoint of effective utilization of limited medical resources. The present study was carried out against this background.

This time, we evaluated the newly developed metastatic site and the deteriorated sites after the first-line TKI therapy in patients with EGFR mutated NSCLC patients. In this study we made four findings. Firstly, we found that metastatic patterns existed at the time of initial diagnosis in EGFR mutated NSCLC patients. Secondly, we found a significant difference in frequency of newly developed distant organ metastasis. Thirdly, there was a clear statistically significant difference in frequency of the exacerbation sites. These results suggest that newly developed metastatic sites did not occur randomly, but had metastatic patterns, and that the presence of metastasis/exacerbation patterns included lesions present at diagnosis. Fourth, in previous studies investigating metastases after first and second or later treatment, metastases were classified into three clusters: lung-liver type, lung-brain type, and bone-liver-other sites type [5]. In this study, which focused only on the metastatic pattern after the first-line TKI, metastases were classified into the following four clusters: bone-other site type, lung-other site type, brain-other site type, and pleura-lung type. In both results of cluster analyses, lung, bone, and brain had a high frequency of distant metastasis. In the present study, the presence of patients with metastases remaining in the thorax was particularly noted. Based on these results, cluster analysis of metastatic patterns might also contribute to personalized medicine.

At the time of initial diagnosis, approximately half of NSCLC patients have distant metastases [12, 13]. Distant metastases in NSCLC patients have been evaluated in many studies [14–16]. These studies have shown that the most common distant metastatic sites were lung, bone, brain, liver and adrenal gland [14–16]. Many patients with distant metastasis had several organ metastases, not single organ metastasis [13–22]. With regard to metastatic sites, there also have been many studies on metastatic sites at autopsy. Metastases found at autopsy were clinically undetectable and could include those that were not clinically relevant [23]. It could be important to diagnose "clinically" meaningful metastases. Knowledge of 'metastasis patterns' could improve the efficiency of detection of metastatic sites by imaging, enable effective treatment of metastatic sites, make better use of medical resources and reduce medical costs. Patients with NSCLC need to be prepared for metastatic disease that can reduce their quality of life, such as metastases to the brain and bones. It seems that patients might develop metastasis to several sites at the same time or show similar metastatic patterns. Therefore, it is desirable to develop an evaluation method regarding metastasis in these patients. Statistical methods for recognizing and analyzing metastases as a pattern has not been established. This study has shown that cluster analysis might be useful for the analysis of metastatic patterns that appear after a particular treatment.

This study provided new information, but with several limitations. First, there was no pathological confirmation of distant metastases determined by diagnostic imaging. Second, although our study used a series of pathologically proven patients with EGFR mutant NSCLC, our small group

of patients might not reflect the overall patient population of the community. Third, this was a retrospective analysis of metastatic data, including cases treated with several EGFR-TKIs. Results might differ if there are differences in the site of recurrence due to different TKIs. Fourth, this study did not focus on the biological mechanism or microscopic evaluation of distant metastases. Despite these limitations, we evaluated the "clinically" meaningful metastases found in currently available images in patients with clinical symptoms. The data obtained from our approach may contribute to the development of new research fields.

Newly developed metastatic and metastatic/exacerbated patterns might be present after the first-line EGFR-TKI therapy in patients with EGFR mutant NSCLC. These metastatic patterns revealed by statistical analysis suggest that the progression of distant metastases includes more than the "mechanical theory" [24] and the "soil seed hypothesis" [25]. Increasing knowledge about specific metastatic patterns can help improve individualized treatment.

*The authors declare no conflict of interest.*

**References**


1. Ball GH, Hall DJ. A clustering technique for summarizing multivariate data. Behav Sci 1967; 12: 153-155.
2. Everitt BS. Cluster analysis. 3rd ed. NY: John Wiley, New York 1993.
3. Pimenta FC, Ribeiro-Dias F, Brandileone MC, et al. Genetic diversity of PspA types among nasopharyngeal isolates collected during an ongoing surveillance study of children in Brazil. J Clin Microbiol 2006; 44: 2838-2843.
4. Kaneko Y, Masuko H, Sakamoto T, et al. Asthma phenotypes in Japanese adults – their associations with the CCL5 and ADRB2 genotypes. Allergol Int 2013; 62: 113-121.
5. Watanabe H, Okauchi S, Yamada H, et al. Application of cluster analysis to distant metastases from lung cancer. Anticancer Res 2020; 40: 413-419.
6. Jin R, Zhao J, Xia L, et al. Application of immune checkpoint inhibitors in EGFR-mutant non-small-cell lung cancer: from bed to bench. Ther Adv Med Oncol 2020 Jun 9; 12: 1758835920930333.
7. Sivars L, Landin D, Grün N, et al. Validation of human papillomavirus as a favourable prognostic marker and analysis of CD8+ tumour-infiltrating lymphocytes and other biomarkers in cancer of unknown primary in the head and neck region. Anticancer Res 2017; 37: 665-673.
8. DO Carmo NG, Sakamoto LH, Pogue R, et al. Altered expression of PRKX, WNT3 and WNT16 in human nodular basal cell carcinoma. Anticancer Res 2016; 36: 4545-4551.
9. Okauchi S, Watanabe H, Yamada H, et al. The prognosis of lung cancer with different metastatic patterns. Anticancer Res 2020; 40: 421-426.
10. Shah R, Girard N, Nagar SP, et al. European and US real-world treatment patterns in patients with epidermal growth factor receptor mutation-positive non-small cell lung cancer: a retrospective medical record review. Drugs Real World Outcomes 2021; 8: 537-545.
11. Tsubata Y, Tanino R, Isobe T. Current therapeutic strategies and prospects for EGFR mutation-positive lung cancer based on the mechanisms underlying drug resistance. Cells 2021; 10: 3192.
12. Satoh H, Kurishima K, Nakamura R, et al. Lung cancer in patients aged 80 years and over. Lung Cancer 2009; 65: 112-118.
13. Kobrinsky NL, Klug MG, Hokanson PJ, Sjolander DE, Burd L. Impact of smoking on cancer stage at diagnosis. J Clin Oncol 2003; 21: 907-913.
14. Satoh H, Ishikawa H, Kamma H, et al. Serum sialyl Lewis X-i antigen levels in non-small cell lung cancer: correlation with distant metastasis and survival. Clin Cancer Res 1997; 3: 495-499.
15. Olak J. Surgical strategies for metastatic lung cancer. Surg Oncol Clin N Am 1999; 8: 245-257.
16. Pfannschmidt J, Dienemann H: Surgical treatment of oligometastatic non-small cell lung cancer. Lung Cancer 2010; 69: 251-258.
17. Oikawa A, Takahashi H, Ishikawa H, et al. Application of conditional probability analysis to distant metastases from lung cancer. Oncol Lett 2012; 3: 629-634.
18. Tamura T, Kurishima K, Watanabe H, et al. Characteristics of clinical N0 metastatic non-small cell lung cancer. Lung Cancer 2015; 89: 71-75.
19. Tamura T, Kurishima K, Nakazawa K, et al. Specific organ metastases and survival in metastatic non-small-cell lung cancer. Mol Clin Oncol 2015; 3: 217-221.
20. Nakazawa K, Kurishima K, Tamura T, et al. Specific organ metastases and survival in small cell lung cancer. Oncol Lett 2012; 4: 617-620.
21. Ishikawa H, Satoh H, Kurishima K, et al. Lung cancer with synchronous brain and bone metastasis. Clin Oncol (R Coll Radiol) 2000; 212: 136-137.
22. Kagohashi K, Satoh H, Ishikawa H, Sekizawa K. Synchronous lung and bone metastasis in lung cancer. Arch Oncol 2003; 11: 297.
23. de Pangher Manzini V, Revignas MG, Brollo A. Diagnosis of malignant tumor: comparison between clinical and autopsy diagnoses. Hum Pathol 1995; 26: 280-283.
24. Ewing J. Neoplastic diseases: a treatise on tumours. 3rd ed. PA: W.B. Saunders, Philadelphia 1928, 76-88.
25. Paget S. The distribution of secondary growths in cancer of the breast. Cancer Metastasis Rev 1989; 8: 98-101.


**Address for correspondence**

Prof. **Hiroaki Satoh**
Mito Medical Center
University of Tsukuba
Tsukuba, Japan
e-mail: hirosato@md.tsukuba.ac.jp

**Submitted:** 15.05.2022
**Accepted:** 10.06.2022