

Table 1 Histopathological distribution of advanced HL among 167 patients

	Number of patients (%)
Nodular lymphocytic predominance	2 (1.2)
Nodular sclerosis	115 (68.9)
Lymphocyte-rich	3 (1.8)
Mixed cellularity	34 (20.4)
Lymphocyte depletion	7 (4.2)
Unclassifiable	6 (3.6)

for advanced HL (128 in JCOG9305 and 72 in JCOG9705), histopathological specimens from 181 patients were available and reviewed, and a consensus diagnosis of HL was reached in 167 (92.3%) (107 in JCOG9305 and 60 in JCOG9705) according to the WHO classification. The remaining 14 patients were diagnosed with diffuse large B cell lymphoma ($n = 4$), T cell-rich B cell lymphoma ($n = 4$), anaplastic large cell lymphoma ($n = 1$), angioimmunoblastic T cell lymphoma ($n = 1$) or other ($n = 4$). The histopathological distribution of the 167 patients with HL is shown in Table 1. Among the HLs, nodular sclerosis ($n = 115$) comprised 68.9% of the whole HL and mixed cellularity ($n = 34$; 20.4%) was the next most frequent subtype in Japan.

3.2 Clinical characteristics

Data from these 167 patients with HL were analyzed. Their clinical characteristics are shown in Table 2. The median age of the patients at diagnosis was 31 years (range 15–69 years). There was a slight male predominance with males comprising 55%. Seventy-six patients (45%) had B symptoms and 49 patients (29%) had extranodal sites.

3.3 OS according to histology and IPS

The 5-year OS of the 167 patients was 88.3% (Fig. 1) (92.3% in JCOG9305 and 81.3% in JCOG9705). The median OS of patients with mixed cellularity was 7.5 years, and those of patients with other histological subtypes was longer than 7.5 years. The 5-year OS of patients with the main histological subtypes was 89.2% in nodular sclerosis and 82.2% in mixed cellularity. The 5-year OS among patients with IPS score of 0, 1, 2, 3, 4, or 5 + 6 was 100% (15 patients), 95.5% (47), 87.5% (40), 86.1% (38), 76.6% (22), or 60.0% (5), respectively (Fig. 2a). Therefore, we failed to identify very high-risk patients by IPS in our study. The OS among patients according to IPS score of 0–2 or 3 and higher is shown in

Table 2 Patient characteristics ($n = 167$)

	Number of patients (%)
Sex	
Male	92 (55.1)
Female	75 (44.9)
Age (years)	
≥ 45	45 (26.9)
< 45	122 (73.1)
Performance status (0/1/2/3)	108/49/7/1
B symptoms	
Yes	76 (45.5)
No	89 (53.3)
Clinical stage	
I/II	83 (49.7)
III	49 (29.3)
IV	35 (21.0)
Bulky mass	
Present	45 (26.9)
Absent	121 (72.5)
Extranodal sites (0/1/ ≥ 2)	106/35/14
Sites of organ involvement	
Liver (yes)	11 (6.6)
Lung (yes)	16 (9.6)
Bone marrow (yes)	10 (6.0)
Other (yes)	29 (17.4)
Baseline hematological data	
Hemoglobin (< 10.5 g/dl)	28 (16.8)
White blood cells ($\geq 15000/\mu\text{l}$)	25 (15.0)
Lymphocytes ($< 600/\mu\text{l}$ or $< 8\%$)	32 (19.2)
Platelets ($< 100000/\mu\text{l}$)	1 (0.6)
Albumin	
< 4 g/dl	99 (59.3)
≥ 4 g/dl	68 (40.7)
Serum LDH	
Elevated	56 (33.5)
Normal	110 (65.9)
Alkaline phosphatase	
Elevated	74 (44.3)
Normal	92 (55.1)
CRP	
Elevated	128 (76.6)
Normal	29 (17.4)
$\beta 2$ -Microglobulin	
> 2 mg/l	39 (23.3)
≤ 2 mg/l	72 (43.1)

Data on performance status, B symptoms, bulky mass, extranodal sites, serum LDH, alkaline phosphatase, CRP or $\beta 2$ -microglobulin were missing in 2, 2, 1, 12, 1, 1, 10 or 56 patients, respectively

Fig. 1 Kaplan–Meier curves for overall survival among all patients with HL ($n = 167$)

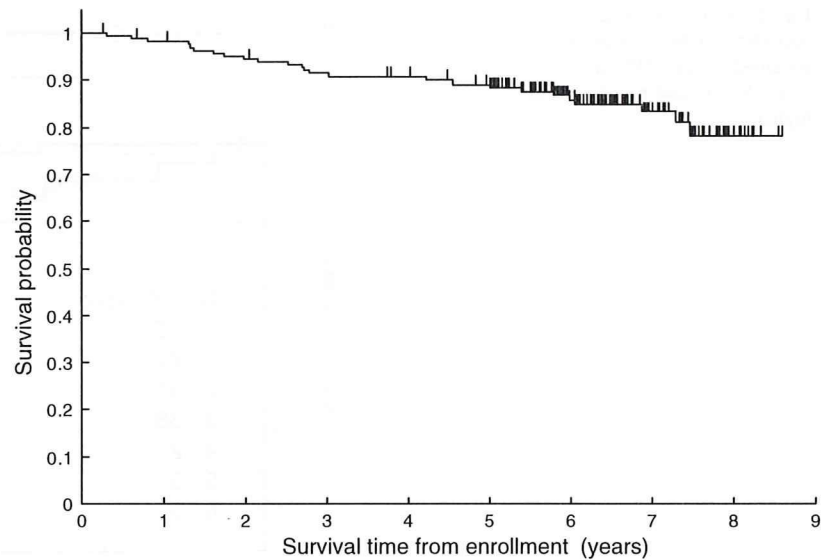


Fig. 2b. An IPS score of 3 or greater was not a significant unfavorable prognostic factor for OS [HR 2.39 (95% CI 1.10–5.21, $p = 0.03$) by univariate analysis and HR 1.20 (95% CI 0.50–2.89, $p = 0.68$) by multivariate analysis with adjustment of other covariates, which were significant in univariate analysis] and thus, IPS was not closely associated with OS. Therefore, we attempted to identify the prognostic factors for OS in Japanese patients with advanced HL by central pathological review.

3.4 Unfavorable prognostic factors by multivariate analysis

Seven unfavorable prognostic factors for OS identified by univariate analysis were male, elevated $\beta 2$ -microglobulin, B symptoms, elevated serum LDH, elevated alkaline phosphatase, clinical stage of III or IV and pathological subtype (mixed cellularity and lymphocyte depletion) (Table 3). Although data of $\beta 2$ -microglobulin were available in only 111 patients (66%), we performed multivariate analysis including $\beta 2$ -microglobulin, but no significant factor was detected. Then, the $\beta 2$ -microglobulin level was excluded from the final multivariate analysis. Male [HR 3.30 (95% CI 1.15–9.52, $p = 0.027$)] and elevated serum LDH [HR 2.41 (95% CI 1.07–5.43, $p = 0.034$)] were significant unfavorable prognostic factors for OS on multivariate analysis (Table 4). Besides, male and elevated serum LDH remained significant in the multivariate analysis including albumin (data not shown). Similarly, elevated serum LDH remained significant in the multivariate analysis including IPS (each 6 categories and 0–2 or 3–6) after sex was excluded (data not shown).

The OS by sex and serum LDH among patients with HL excluding those with unclassifiable histopathology is

shown in Fig. 3a and b, respectively. The 5-year OS was 82.4% in males and 94.4% in females, and 82.4% in patients with elevated serum LDH and 90.5% in patients with normal serum LDH.

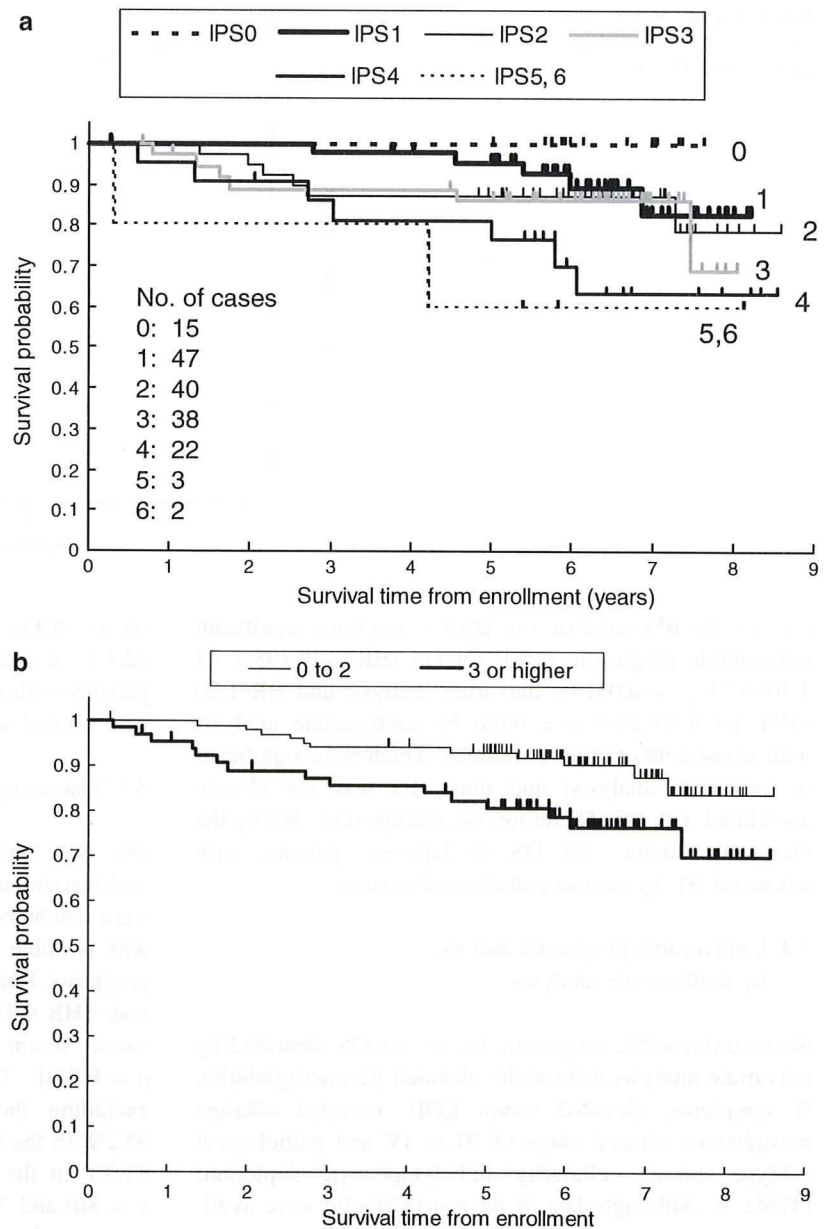
3.5 Risk group model

The two important prognostic factors identified by the multivariate analysis, i.e., male and elevated serum LDH, were combined in a prognostic index to create risk groups with possible values of 0, 1 and 2 in order of worsening prognosis. Hazard ratios of the final model were as follows: male [HR 4.91 (95% CI 1.84–13.13, $p = 0.002$)] and elevated serum LDH [HR 2.72 (95% CI 1.25–5.89, $p = 0.01$)]. The 5-year OS among patients with HL excluding those with unclassified histopathology was 95.2% in the low-risk group (no adverse factor, $n = 47$), 87.9% in the intermediate-risk group (1 adverse factor, $n = 86$) and 73.3% in the high-risk group (2 adverse factors, $n = 27$). Data on serum LDH were missing in 1 patient. The OS curves of the 3 risk groups are shown in Fig. 4 (corrected $p = 0.004$ by fivefold cross-validation).

4 Discussion

It is recognized that there is an uneven geographical distribution of malignant lymphomas throughout the world. Namely, the incidence of T cell lymphoma is relatively high in Asia compared with Western countries. On the contrary, the incidence of HL in Japan was reported to be 4.4% of malignant lymphomas and this is relatively low compared with those in Western countries [11, 15, 16]. The low incidence of HL in Japan limited the evaluation of the

Fig. 2 Overall survival according to the international prognostic score (IPS) **a** 1, 2, 3, 4 and 5 + 6, and **b** 0–2 or 3 or higher



applicability of IPS or other prognostic factors to Japanese patients with advanced HL treated with an established protocol considered to be state-of-the-art combination chemotherapy or chemo-radiotherapy. To our knowledge, this is the first report to validate the IPS comprehensively and to analyze the conventional prognostic factors for OS in a large number of Japanese patients with advanced HL treated with established protocols of state-of-the-art combination chemotherapy or chemo-radiotherapy and diagnosed by central pathological review. The histopathological distributions of advanced HL according to the WHO classification in 167 patients in Japan were determined, and showed that the proportion of patients with

nodular sclerosis in Japan (68.9%) was higher, while the proportion with mixed cellularity (20%) was similar to those in Western countries [12, 16]. The survival of each HL subtype in Japan was similar to those in Western countries [12].

In our study, there were only 5 patients with IPS score of 5 or higher, accounting for only 3% of the entire study population. Their 6-year OS was 60%, indicating that a distinct group of patients at very high risk could not be determined on the basis of IPS. Even in the original IPS paper, only 7% of the patients had a score of 5 or higher representing a very high risk and had a 59% of OS at 5 years. The results were very similar to our study.

Table 3 Univariate survival analysis

		Hazard ratio	95% CI	<i>p</i>
Sex	Male	4.30	(1.62–11.45)	0.004
β 2-Microglobulin	>2 mg/l	4.78	(1.49–15.27)	0.008
B symptoms	Yes	0.31	(0.13–0.74)	0.009
Serum LDH	Elevated	2.55	(1.18–5.51)	0.02
Alkaline phosphatase	Elevated	2.53	(1.13–5.67)	0.03
Clinical stage	III/IV	2.41	(1.05–5.56)	0.04
Histopathology	MC&LD	2.21	(1.02–4.83)	0.05
Albumin	<4 g/dl	2.44	(0.98–6.09)	0.06
Age (years)	\geq 45	1.80	(0.82–3.96)	0.15
Hemoglobin	<10.5 g/dl	0.36	(0.08–1.51)	0.16
White blood cells	\geq 15000/ μ l	1.83	(0.73–4.55)	0.20
Lymphocytes	<600/ μ l or <8%	1.63	(0.69–3.88)	0.27
Clinical stage	IV	1.40	(0.59–3.32)	0.45
Extranodal sites	Yes	0.97	(0.40–2.39)	0.95

Table 4 Multivariate survival analysis

		Hazard ratio	95% CI	<i>p</i>
Sex	Male	3.30	(1.15–9.52)	0.03
Serum LDH	Elevated	2.41	(1.07–5.43)	0.03
B symptoms	Yes	2.26	(0.85–6.06)	0.10
Alkaline phosphatase	Elevated	1.94	(0.70–5.37)	0.21
Histopathology	MC&LD	1.73	(0.75–4.00)	0.20
Clinical stage	III/IV	0.87	(0.31–2.47)	0.80

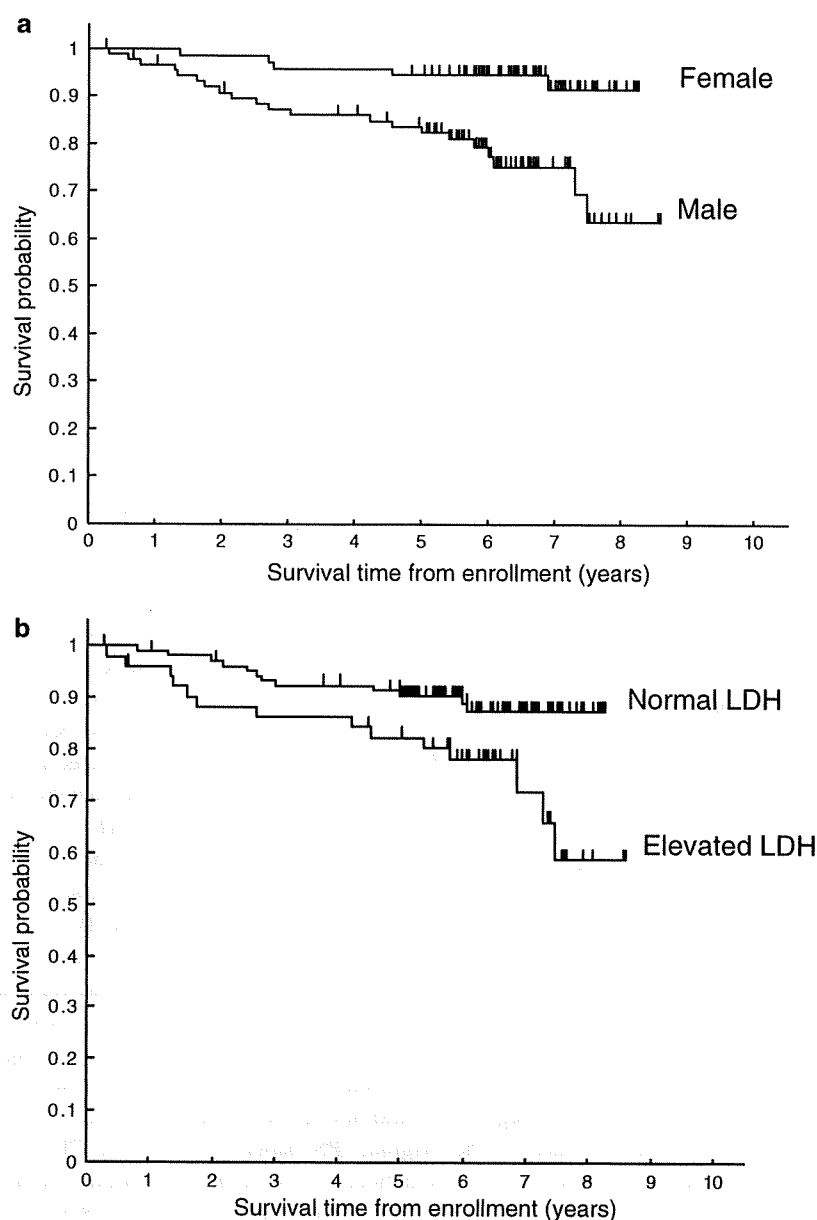
Therefore, in our study, it can be concluded that a distinct group of patients at a very high risk could not be identified by IPS, as same as stated in the original IPS paper [7]. Furthermore, the survival curves of patients with an IPS score of 1, 2 or 3 were not clearly separated from each other (Fig. 2). Therefore, IPS was not closely associated with OS in our study. It has been reported that the IPS score 0–4 versus 5 or 6 was found to have prognostic significance for disease-specific survival in a report of large number of Japanese patients with HL treated variously, in which the presence of T cell and/or cytotoxic antigen in Hodgkin's and Reed-Sternberg cells also showed a significant poor prognosis [17]. However, in that study, neither patient number nor survival rate of patients in each IPS score was shown at all, thus it may be said that IPS was not adequately validated in Japanese patients with advanced HL treated with state-of-the-art combination chemotherapy or chemo-radiotherapy. As it had been concluded that a distinct group of patients at very high risk could not be identified by the IPS [7], attempts have been made to

determine more suitable factors that could detect the poor-risk population among patients with HL [17–19]. Namely, it was reported that the number of involved anatomic sites combined with the IPS [18] or interleukin-10 (IL-10) level added to the IPS [19] could detect the subgroup of HL patients with poor prognosis.

As the initial treatments for HL led to excellent outcomes and the rescue treatments could improve the clinical outcomes, the prognostic factors for OS might be more important than the prognostic factors for progression-free survival (PFS) in patients with advanced HL. In our study in which the diagnosis of HL was based on central pathological review, OS was independently affected by male and elevated serum LDH on multivariate analysis. Only male and clinical stage of III or IV among the 7 factors in the IPS were significantly associated with poor OS in the univariate analysis, and male remained significant in the multivariate analysis. The German Hodgkin Study Group suggested that hematotoxicities were more pronounced in females although this did not translate into increased infection, and female patients had similar response rates as males but fewer relapses and deaths, leading to a significantly better freedom from treatment failure in a large retrospective analysis [20]. Sex might be associated with the metabolism of anticancer drugs [21]. Elevated serum LDH was previously reported to be prognostically unfavorable in advanced HL [22, 23] and is also one of the most important factors in the international prognostic index of non-HL [24]. Therefore, elevated serum LDH might reflect the total status of HL, including both constitutional and disease-related elements.

As a post hoc sensitivity analysis, we also performed the analysis using stepwise variable selection methods, and the results were shown that male [HR 6.18 (95% CI 2.28–16.70, $p < 0.001$)] and elevated serum LDH [HR 2.87 (95% CI 1.32–6.24, $p = 0.008$)] remained significant and serum albumin level of less than 4 g/dl [HR 3.38 (95% CI 1.35–8.51, $p = 0.01$)] was also significant. Although serum albumin was significantly correlated with B symptom ($p < 0.001$) and serum alkaline phosphatase ($p = 0.001$), serum albumin did not show the significance over male and serum LDH as the prognostic factor and did not show the prognostic relevance for OS in univariate analyses. Although stepwise variable selection method was widely used, model by stepwise method is not necessarily considered the best with regard to the statistical issues, which were often discussed and criticized [25–28]. In this study, analysis was performed following the prospectively planned method, and the final model was evaluated by cross-validation, one of the internal validation methods to resolve these statistical issues. The OS curves between risk groups derived from our final model were significantly different, and the results were validated by cross-validation. Based

Fig. 3 Overall survival according to **a** sex ($n = 167$) and **b** serum LDH ($n = 166$)

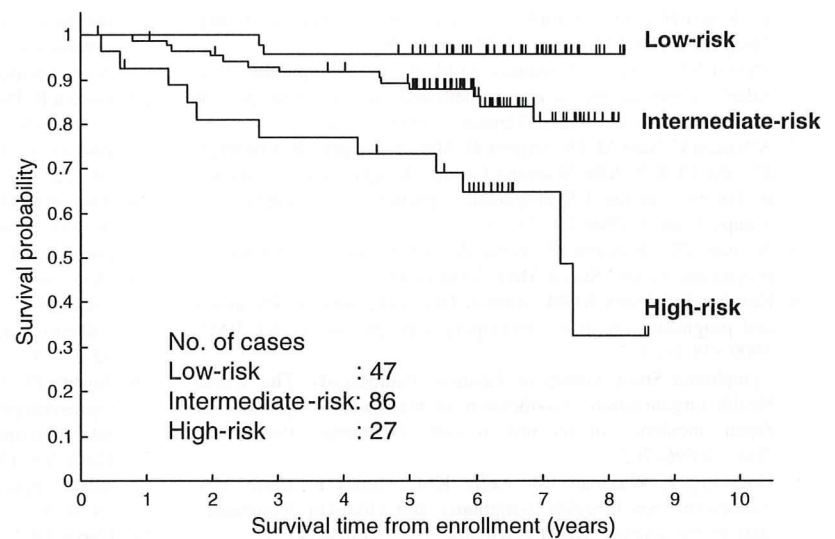


on a combination of model fit and parsimony considerations, our final model incorporated two prognostic factors: male and elevated serum LDH.

Generally, complex models with a large number of prognostic predictors are not practical and simple models are easier to evaluate and are preferable in routine clinical practice. The prognostic model for advanced HL in our study, including sex and serum LDH, was considered to be very simple. However, it is discussed that prognostication with the prospective studies has a limitation by the exclusion of patients with poor condition and prognostic models using data of prospective study might be difficult for generalization. On this concern, adequate consideration should be necessary.

Prognosis of the patients with advanced HL is improved in advance of treatment, and then prognostic factors may differ according to the state of the treatment. In the original IPS paper, eligible patients with advanced HL for the original IPS study were limited to those who were 15–65 years old and were treated with an established protocol still considered to be state-of-the-art, with at least four planned cycles of combination chemotherapy (preferably containing doxorubicin) with or without radiotherapy. This means that IPS was established in the patients who could be safely treated with state-of-the-art therapy, excluding both elderly patients of more than 65 years and those who were poorly treated probably because of poor condition. Nonetheless, IPS has been used widely, because

Fig. 4 Overall survival among patients with HL excluding those with unclassified histopathology according to the number of unfavorable prognostic factors (male and elevated serum LDH). Low risk, intermediate risk and high risk indicate 0, 1 and 2 risk factors, respectively. Data on serum LDH were not available in one patient, and this patient was excluded from this analysis



everybody wants to know the prognostic state of patients with advanced HL treated with state-of-the-art therapy. In our study, eligible patients for the prognostic analysis are almost same as patients for the original IPS study. Then, our prognostic model could be accepted for general use, although further studies should be warranted to validate our prognostic model.

Unfortunately, data on the $\beta 2$ -microglobulin level of 56 patients were missing. In univariate analysis, $\beta 2$ -microglobulin was found to be highly significant in 111 patients. However, multivariate analysis including $\beta 2$ -microglobulin revealed that there was no significant factor detected. Then, $\beta 2$ -microglobulin was excluded from the final multivariate analysis for OS. Serum $\beta 2$ -microglobulin levels are known to reflect renal function and membrane turnover, the latter of which is associated with tumor mass and growth rate. Elevated $\beta 2$ -microglobulin level was reported to predict poor survival in several hematological malignancies including low-grade lymphoma [29], large cell lymphoma [30] and HL [31–33]. Interestingly, Vassilakopoulos et al. [33] reported that the $\beta 2$ -microglobulin level was a powerful independent prognostic factor for OS, but not for failure-free survival in optimally treated patients with HL. The prognostic impact of $\beta 2$ -microglobulin on OS should be re-evaluated in future.

In conclusion, despite the limitation of a small number of patients, our prognostic model was considered to be a simple method of predicting OS in Japanese patients with advanced HL. Further studies to validate our prognostic model and to re-evaluate the prognostic impact on OS of sex and serum LDH combined with $\beta 2$ -microglobulin are warranted.

Acknowledgments We thank Kiyoshi Mukai (Tokyo Medical University), Shigeo Nakamura (Nagoya University), and Kouichi Ohshima (Kurume University) for pathological review as members of

an expert panel. This study was supported by Grants-in-Aid for Cancer Research from the Ministry of Health, Labour and Welfare of Japan (2S-1, 5S-1, 8S-1, 11S-1, 11S-4, 14S-1, 14S-4, 17S-1, 17S-5).

References

1. Evens AM, Hutchings M, Diehl V. Treatment of Hodgkin lymphoma: the past, present, and future. *Nat Clin Pract Oncol*. 2008;5:543–56.
2. Franklin J, Pluetschow A, Paus M, Specht L, Anselmo AP, Aviles A, et al. Second malignancy risk associated with treatment of Hodgkin's lymphoma: meta-analysis of the randomised trials. *Ann Oncol*. 2006;17:1749–60.
3. Schonfeld SJ, Gilbert ES, Dores GM, Lynch CF, Hodgson DC, Hall P, et al. Acute myeloid leukemia following Hodgkin lymphoma: a population-based study of 35 511 patients. *J Natl Cancer Inst*. 2006;98:215–8.
4. Hodgson DC. Hodgkin lymphoma: the follow-up of long-term survivors. *Hematol Oncol Clin N Am*. 2008;22:233–44.
5. De Bruin ML, Dorresteijn LDA, van't Veer MB, Krol ADG, van der Pal HJ, Kappelle AC, et al. Increased risk of stroke and transient ischemic attack in 5-year survivors of Hodgkin lymphoma. *J Natl Cancer Inst*. 2009;101:928–37.
6. Zander T, Wiedemann S, Wolf J. Prognostic factors in Hodgkin's lymphoma. *Ann Oncol*. 2002;13(Suppl 1):67–74.
7. Hasenclever D, Diehl V, for the international prognostic factors project on advanced Hodgkin's disease. A prognostic score for advanced Hodgkin's disease. *N Engl J Med*. 1998;339:1506–14.
8. Ogura M, Kagami Y, Itoh K, Sasaki Y, Kinoshita T, Tobinai K, et al. Phase II study of ABVD therapy to advanced stage Hodgkin's disease in Japan: JCOG study 9305. *Proc Am Soc Clin Oncol*. 2001;230a.
9. Ogura M, Morishima Y, Itoh K, Tobinai K, Kinoshita T, Okamoto M, et al. Dacarbazine (DTIC) cannot be deleted from ABVD therapy for advanced Hodgkin lymphoma (HL): Japan Clinical Oncology Group (JCOG) Study 9705. *Proc Am Soc Clin Oncol*. 2003;573a.
10. Kinoshita T, Hotta T, Tobinai K, Kobayashi T, Ishizuka N, Tomonaga M, et al. Lymphoma Study Group of the Japan Clinical Oncology Group (JCOG-LSG): a randomized controlled trial investigating the survival benefit of dose-intensified multidrug combination chemotherapy (LSG9) for intermediate- or high-

- grade non-Hodgkin's lymphoma: Japan Clinical Oncology Group Study 9002. *Int J Hematol.* 2004;80:341–50.
11. Proctor SJ, Wilkinson J, Sieniawski M. Hodgkin lymphoma in the elderly: a clinical review of treatment and outcome, past, present and future. *Crit Rev Oncol/Hematol.* 2009;71:222–32.
 12. Allemani C, Sant M, De Angelis R, Marcos-Gragera R, Coebergh JW, the EURO CARE Working Group. Hodgkin disease survival in Europe and the US prognostic significance of morphologic groups. *Cancer.* 2006;107:352–60.
 13. Altman DG, Royston P. What do we mean by validating a prognostic model? *Statist Med.* 2000;19:453–73.
 14. Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. *BMJ.* 2009;338:1373–7.
 15. Lymphoma Study Group of Japanese Pathologists. The World Health Organization classification of malignant lymphomas in Japan: incidence of recently recognized entities. *Pathol Int.* 2000;50:696–702.
 16. Hochberg J, Waxman IM, Kelly KM, Morris E, Cairo MS. Adolescent non-Hodgkin lymphoma and Hodgkin lymphoma: state of the science. *Brit J Haematol.* 2009;144:24–40.
 17. Asano N, Oshiro A, Matsuo K, Kagami Y, Ishida F, Suzuki R, et al. Prognostic significance of T-cell or cytotoxic molecules phenotype in classical Hodgkin's lymphoma: a clinicopathologic study. *J Clin Oncol.* 2006;24:4626–33.
 18. Vassilakopoulos TP, Angelopoulou MK, Siakantaris MP, Kontopidou FN, Dimopoulou MN, Barbounis A, et al. Prognostic factors in advanced stage Hodgkin's lymphoma; the significance of the number of involved anatomic sites. *Euro J Hematol.* 2001;67:279–88.
 19. Axedorph U, Sjoberg J, Grimfors G, Landgren O, Porwit-MacDonald A, Bjorkholm M. Biological markers may add to prediction of outcome achieved by the international prognostic score in Hodgkin's disease. *Ann Oncol.* 2000;11:1405–11.
 20. Klimm B, Reineke T, Haverkamp H, Behringer K, Eich HT, Josting A, et al. Role of hematotoxicity and sex in patients with Hodgkin's lymphoma: an analysis from the German Hodgkin study group. *J Clin Oncol.* 2005;23:8003–11.
 21. Dobbs NA, Twelwvs CJ, Gillies H, James CA, Harper PG, Rubens RD. Gender affects doxorubicin pharmacokinetics in patients with normal liver biochemistry. *Cancer Chemother Pharmacol.* 1995;36:473–6.
 22. Straus DJ, Gaynor JJ, Myers J, Merke DP, Caravelli J, Chapman D, et al. Prognostic factors among 185 adults with newly diagnosed advanced Hodgkin's disease treated with alternating potentially noncross-resistant chemotherapy and intermediate-dose radiation therapy. *J Clin Oncol.* 1990;8:1173–86.
 23. García R, Hernández JM, Caballero MD, González M, Galende J, del Cañizo MC, et al. Serum lactate dehydrogenase level as a prognostic factor in Hodgkin's disease. *Br J Cancer.* 1993;68:1227–31.
 24. The international non-Hodgkin's lymphoma prognostic factors project. A predictive model for aggressive non-Hodgkin's lymphoma. *N Engl J Med.* 1993;329:987–94.
 25. Derkson S, Keselman HJ. Backward, forward and stepwise automated subset selection algorithms: frequency of obtaining authentic and noise variables. *Br J Math Stat Psychol.* 1992;45:265–82.
 26. Austina PC, Tu JV. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. *J Clin Epidemiol.* 2004;57:1138–46.
 27. Flack VF, Chang PC. Frequency of selecting noise variables in subset regression analysis: a simulation study. *Am Stat.* 1987;14:84–6.
 28. Copas JB, Long T. Estimating the residual variance in orthogonal regression with variable selection. *Statistician.* 1991;40:51–9.
 29. Litam P, Swan S, Cabanillas F, Tucker SL, McLaughlin P, Hagemester FB, et al. Prognostic value of serum $\beta 2$ microglobulin in low-grade lymphoma. *Ann Intern Med.* 1991;114:855–60.
 30. Swan F, Velasquez WS, Tucker S, Redman JR, Rodriguez MA, McLaughlin P, et al. A new serologic staging system for large-cell lymphomas based on initial $\beta 2$ -microglobulin and lactate dehydrogenase levels. *J Clin Oncol.* 1989;7:1518–27.
 31. Dimopoulos MA, Cabanillas F, Lee JJ, Swan F, Fuller L, Allen PK, et al. Prognostic role of serum $\beta 2$ microglobulin in Hodgkin's disease. *J Clin Oncol.* 1993;11:1108–11.
 32. Chronowski GM, Wilder RB, Tucker SL, Ha CS, Sarris AH, Hagemester FB, et al. An elevated serum $\beta 2$ -microglobulin level is an adverse prognostic factor for overall survival in patients with early-stage Hodgkin disease. *Cancer.* 2002;95:2534–8.
 33. Vassilakopoulos TP, Nadali G, Angelopoulou MK, Siakantaris MP, Dimopoulou MN, Kontopidou FN, et al. The prognostic significance of $\beta 2$ -microglobulin in patients with Hodgkin's lymphoma. *Haematologica.* 2002;87:701–8.

分子標的治療薬の評価

Clinical trial designs for evaluation of target-based drugs

山本精一郎

Key words : 分子標的薬, 研究デザイン, 個別化治療, 予後因子, 予測因子

1. 分子標的治療薬の評価とは

分子標的治療薬を臨床試験により評価すること、すなわち分子標的治療薬の治療開発の研究デザインはこれまでの細胞障害性薬剤の治療開発と大きく変わるものではない。あえて挙げるとすれば、分子標的薬は、もともと特定の治療ターゲットの存在を想定して開発された薬剤であるので、治療開発の過程で治療効果が現れる集団を同定しながら治療開発を進める必要があるということであろう。分子標的薬の場合、対象を特定せずに治療開発を行い、広い対象に対して標準治療との比較試験を行うと、効果のない対象を含んでしまったために標準治療との間で治療効果の差を検出できず、特定の対象に対して高い効果のある薬剤の開発を中止してしまうということが起こる可能性がある。別の言い方をすると、効果のある対象でなく、効果のない対象をうまく除外することによって薬剤の力を適切に評価することが重要であるといえる。しかしながら、効果のある対象を除外しないことも重要である。

これは、これまでの細胞障害性薬剤においても、特に治療効果が高いサブグループを探索する、治療効果のないサブグループを同定する、という試みの中で行われてきたことである。そこで本稿では、治療効果の異なるサブグループ

を探索するという観点から、これまで実施されてきた、あるいは現在計画中の臨床試験を例にしながら、分子標的薬評価のための臨床試験デザインの立て方について考えてみたい。換言すれば、個別化治療開発の臨床試験デザインについて、著者の整理を紹介することを本稿の目的としたい。

2. 個別化治療評価にはランダム化比較試験が必要

治療効果の異なる集団を同定するためには、その集団を特定するマーカーが必要となる。このようなマーカーを予測因子(または治療効果予測因子)と呼ぶ。すなわち、予測因子となるマーカーを探すことが個別化治療を進めることであるともいえる。統計的にはこれを‘予後に関して治療とマーカーの間に交互作用あり’という。これに対して、治療を受けていない場合に予後を予測するマーカーを予後因子と呼ぶ。このように予後因子、予測因子を明確に区別することは、個別化治療開発を考えるうえでの理解の助けとなる。

図1に予後因子、予測因子の別を示す。マーカーが予測因子であるためには、少なくともどこかの群では治療効果があることを示す必要がある。これは、少なくともどこかの群でランダム化比較試験を行う必要があることを示し

Seiichiro Yamamoto: Cancer Information Services and Surveillance Division, Center for Cancer Control and Information Services, National Cancer Center 国立がんセンターがん対策情報センター がん情報・統計部

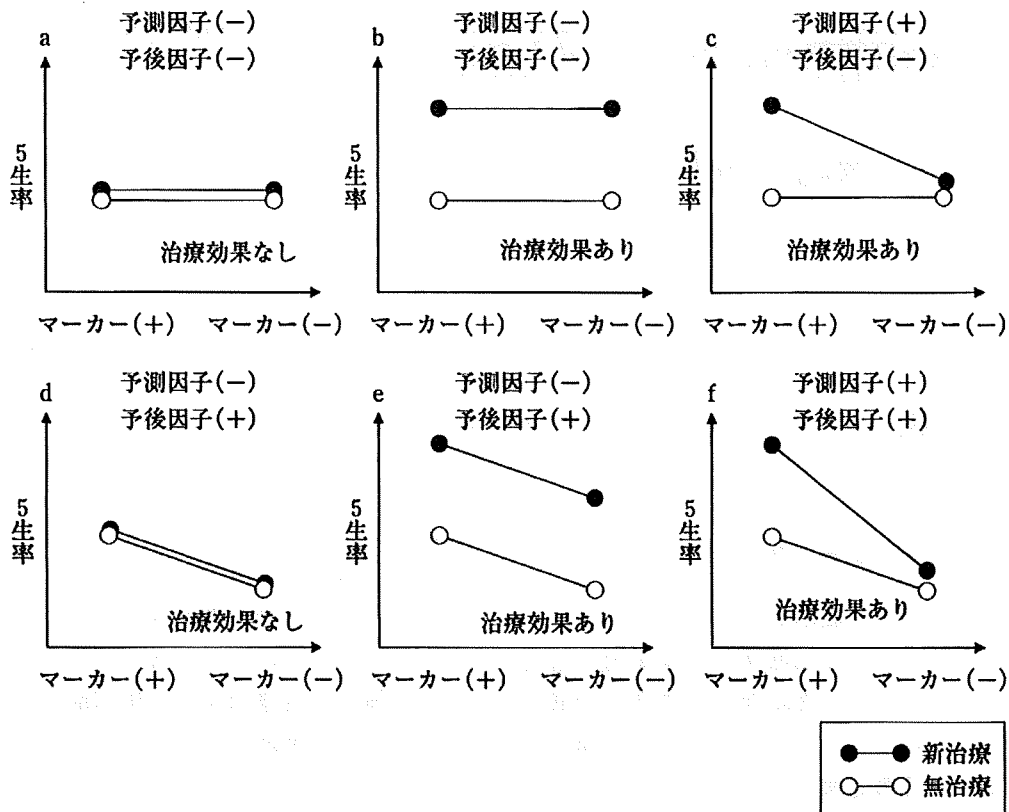


図1 予測因子と予後因子

ている。

3. 個別化治療開発のためのランダム化比較試験デザイン

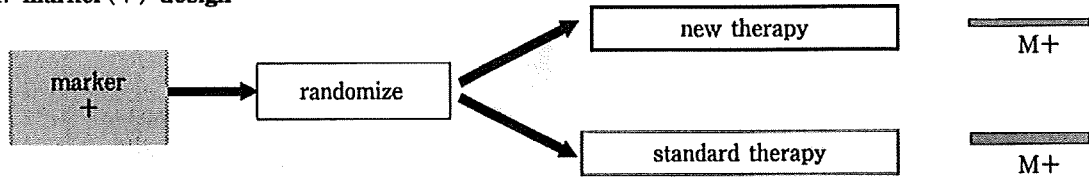
a. Marker(+) design

予測因子を探すという考えに基づく個別化治療開発に対して、大きく分けて3種類の臨床試験デザインが提案されている(図2)¹⁾。まず一つ目は'marker(+) design'といえるもので、マーカーを測定し、(+)の人だけをランダム化して新治療と標準治療を行うものである。いわゆる enrichment design もこのデザインの一つといえる。大きな効果を予測できるマーカー(+)の集団のみ臨床試験を行うような enrichment の場合には試験を行う賛同は得られやすいかもしれない。enrichment の場合、マーカー(-)の集団には治療効果がないことを想定しているか、あっても小さいであろうことを想定している。しかし、マーカー(+)の集団で治療効果が観察されても、マーカー(-)に対しては情

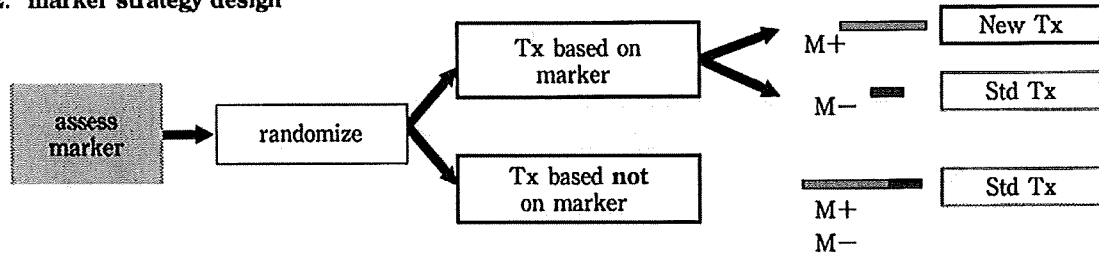
報が得られないので、enrichment だけではマーカーが予測因子かどうか分からないし、もしマーカー(-)のサブグループで治療効果がある場合には、その対象の患者に有効な治療を提供できないことになる。それを解消するために、最初の比較試験では enrich した対象で比較し、治療効果が証明されればマーカー(-)の集団で開発を行うというストラテジーも考えられるかもしれない。

PottiらはStage IAの非小細胞肺癌に対し、手術時に得られた組織における遺伝子発現を用いてその後の予後を予測する予測モデルを構築した²⁾。それにより、再発高リスク群と低リスク群に分けることができるとし、この予測モデルを用いて新たな臨床試験を行うことを提唱している(図3)(現在実施中とのこと、私信)。同じStage IAの患者でも、彼らのモデルにより分類された低リスク群の予後は十分によいが、高リスク群の予後はかなり悪く、この群に対しての新たな治療開発が必要だと考えられる。したが

1. marker(+) design



2. marker strategy design



3. all comers design

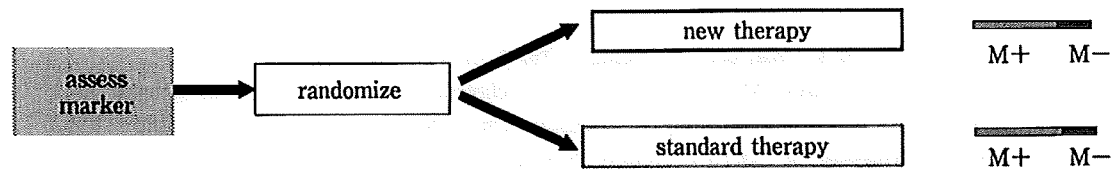
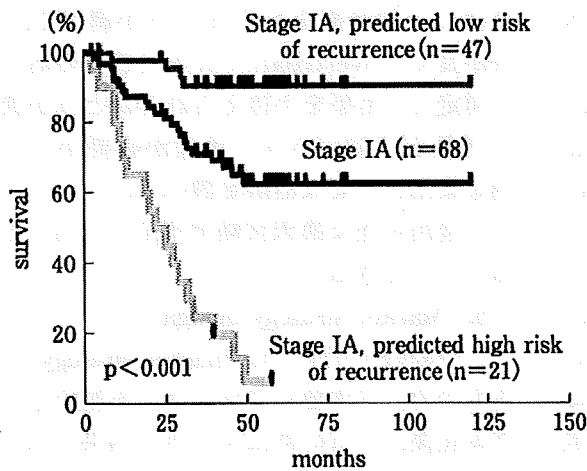


図 2 分子標的薬開発に有効として提案されている臨床試験デザイン



next step

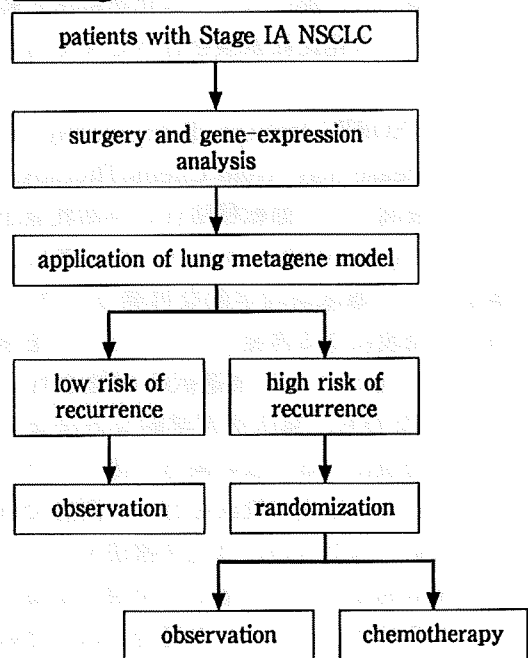


図 3 Marker(+) design の例 1

[Reprinted from Potti A, et al: A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. N Engl J Med 355(6): 570-580, Copyright © (2006) with permission Massachusetts Medical Society. All rights reserved]

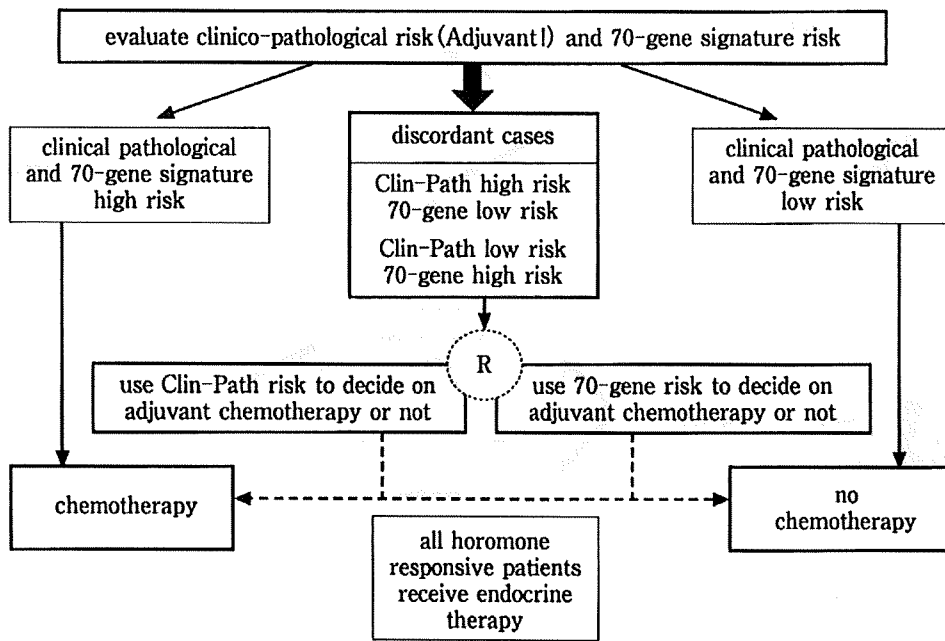


図4 Marker(+) designの例2(EORTC MINDACT Trial)

[Reprinted from Buyse M, et al: Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst 98(17): 1183-1192, 2006 by permission of Oxford University Press]

って、まず手術材料を用いて予後を予測し、低リスクと予測された群は経過観察とし、高リスクと予測された群に対しては化学療法と経過観察のランダム化比較試験を行うとするものである。

また、EORTC MINDACT(Microarray In Node (一) Disease may Avoid ChemoTherapy) Trialでは、腋窩リンパ節転移陰性の早期乳癌患者に対し、Adjuvant!というソフトウェアによって推定された臨床病理学的な再発リスクと70の遺伝子発現による再発リスクのスコアを求め、これらに基づいた治療選択の評価を行っている³⁾。すなわち、臨床病理学的な方法ならびに遺伝子発現による方法の両方で高リスクと推定されたものには化学療法を行い、両方で低リスクと推定されたものには化学療法を行わず、臨床病理学的なリスクと遺伝子発現によるリスクによる予測が不一致だった場合には、その対象者に対し、化学療法ありなしを比較するランダム化比較試験を行うものである(図4)。

これらの例でもわかることは、marker(+) designとは、①効果があると考えられる対象を

選択してその対象に臨床試験を行う(enrichment)、といったものだけでなく、②予後が悪く、更に治療開発が必要な集団を同定し、その集団に追加治療を行うかどうか調べる、③予後が良く、overtreatmentの可能性のある集団を同定し、治療を手控えられようかどうか調べる、④予後が予測できず、治療が必要かどうか不明な集団に必要な治療を調べる、といった場合にも適用可能な臨床試験デザインであると考えることができる。

b. Marker strategy design

二つめのデザインは'marker strategy'といえるもので、まず始めにマーカーを測定し、ランダム化後、一方の群はマーカー結果に基づいて治療を決定し(marker-based)、もう一方の群はマーカーとは関係なく標準治療を行い(not marker-based)、これを比較するというものである。プライマリな解析は、marker-based or notのストラテジーを比較する、つまり試験に参加した対象者全体を比較する。このデザインは、マーカーによって規定されるそれぞれのサブグループに対して異なる治療を行うという

Drug tested

- シスプラチン+ゲムシタビン
- ドキソルビシン
- バクリタキセル
- ミトキサントロン
- ミトキサントロン+バクリタキセル
- トポテカン
- トレオスルファン
- トレオスルファン+ゲムシタビン
- トレオスルファン+エビルピシン
- シスプラチン+エトポシド
- エトポシド

ASCO website
Cree et al. Proc ASCO Vol 23,
No. 16S, Part 1, 2005: 5008

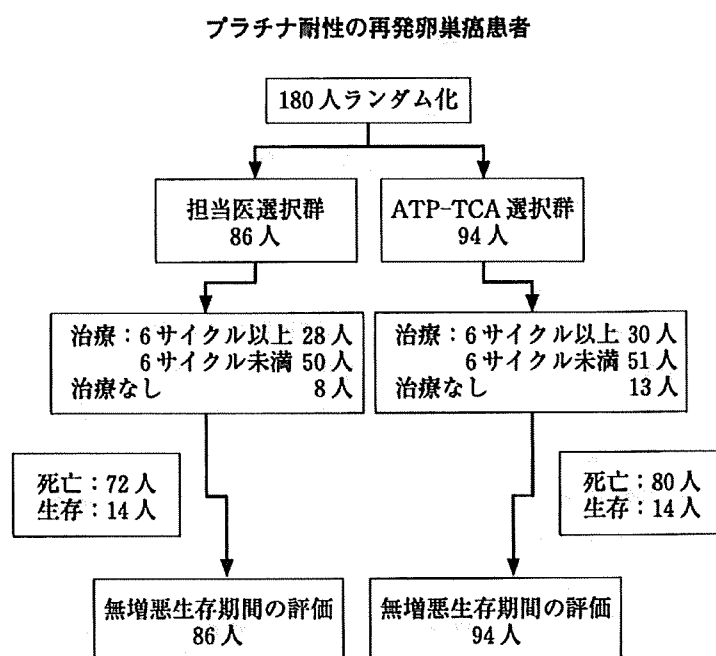


図5 Marker strategy design の例

‘個別化治療戦略そのもの’を評価できるデザインと考えられる。ただ、効果のある集団を同定する=予測因子マーカーを同定する、あるいは proof of principle を調べるための臨床試験デザインではない。なぜなら、マーカー(-)のサブグループには、marker-basedの群でもそうでない群でも同じ治療が行われているため、プライマリな解析で有意に差があった場合でも、マーカー(-)に対する治療効果がわからず、マーカー(+)の群のみで治療効果があるのかどうか(交互作用があるかどうか)調べることが不可能だからである。プライマリな解析(対象者全体の解析)で有意な差がみられなかった場合には、マーカー(+)サブグループでの治療効果も不明である。なぜなら、マーカー(+)サブグループで差があり、マーカー(-)で差がなかった場合、マーカー(+)サブグループの割合が多ければ全体として有意差あり、少なければ有意差なしとなる可能性があり、結果が試験参加者の中のマーカー(+)の割合に依存してしまうからである。

Kurbacherらは、再発卵巣癌患者に対し、ATP-TCA(tumor chemosensitivity assays)の結果に基づいて化学療法を選択したパイロット研

究の結果と経験的に化学療法を選択しているヒストリカルなデータを比較し、ATP-TCAに基づいて化学療法を選択した方がヒストリカルなデータに比べ、奏効割合や無増悪生存期間において良好な成績を示すことを報告した⁹⁾。Creeらは、この結果をもとに、プラチナ耐性の再発卵巣癌患者に対し、13種類の化学療法の中からATP-TCAに基づいて化学療法を選択するという戦略と担当医が化学療法を選択するという戦略を比較するランダム化比較試験を行った⁹⁾。結果として、180人をランダム化して86人が担当医選択群、94人がATP-TCA選択群に分けられ、無増悪生存期間ならびに全生存期間を比較したところ、大きな差はみられないという結果を得た(図5)。この例では、対照群として効果が検証された標準治療がない場合でも、担当医や患者により決定された日常診療に基づく治療に対して、一定の戦略によって治療方法を決定する治療戦略どうしを比較することができ、魅力的である。ただし、結果は対照群の治療結果に依存するため、治療方法選択に大きな違いがなければ大きな差が生じず、治療戦略の効果の差を検出することが難しいことが予想される。

c. All comers design

三つめは 'all comers' と呼ばれるデザインで、まず始めにマーカーを測定するが、その結果によらずランダム化し、一方には新治療、もう一方には標準治療を行うというデザインである。プライマリな解析としてマーカーによらず全対象者の群間比較を行い、セカンダリな解析としてマーカー(+)/(−)それぞれで治療効果の群間比較を行う。このデザインを用いることができれば、マーカー(+)/(−)を合わせた集団に対する治療効果に加え、マーカーサブグループ間で治療効果が異なるかどうか調べられるので、マーカーが治療効果の予測因子であるかどうかを調べることができる。

このデザインは、最初にマーカーを測定しないで、試験終了後(ランダム化後)にマーカーを測定するのと同じではないかと考えられるかもしれない。Fisherらはリンパ節転移陰性、ER(+)の乳癌女性に対し、化学療法の効果を調べる臨床試験を行い、2,306人の乳癌患者をランダムにタモキシフェン群、タモキシフェン+化学療法(MFT)群、タモキシフェン+化学療法(CMFT)群の3群に分け、化学療法群の方が予後が良いことを示した(NSABP B20)⁶⁾。

Paikらは、この試験に参加した対象者の組織と臨床試験の結果を用いて、遺伝子発現に基づく Oncotype DXによる再発スコア(RS)の有効性を調べた⁷⁾。NSABP B20の対象者のうち、651人(227人がタモキシフェン群であり、424人が化学療法群)について Oncotype DXによる再発スコアが調べられた。再発スコアにより高リスク、中リスク、低リスクと分けた場合、高リスク群のみ化学療法群とタモキシフェン群に予後の差が認められた(図6)。この結果により、化学療法の追加により恩恵を受けるのは高リスク群のみである、言い換えると高リスク群のみ化学療法を追加すべきであると主張している。

この研究はまさに事後的な解析によって、all comers design 的な臨床試験を仮想的に行った例と考えられる。もし、事前に再発スコアを測定し、測定した対象者に対しランダム化比較試験を行い、その結果において3つのリスクカ

テゴリで予後を比較し、高リスク群でのみ化学療法による予後が優れていることが示されれば、リスクスコアと治療法間に交互作用があり、高リスク群では化学療法を追加することが推奨されると結論づけることができるであろう。しかしながら、この例のように事後的に測定を行った場合には、ランダム化した全例に再発スコアを測定できなかったためにリスクスコアごとの治療効果比較はランダムな比較とはいえず、差があったとしてもRCTによるエビデンスということとはできない。また、事後にマーカーを測定した場合には、群ごとのサンプルサイズ設計も十分行われているとはいえず、差がない場合にも検出力不足である可能性がある。ただ、それでも、観察研究のデータより、ランダム化比較試験に基づいた事後的な解析の方が結果の妥当性は高いと考えられるのは明らかであろう。

事後的に行った場合のもう一つの問題点として、事前に特定しないマーカーをいろいろと測定し、それぞれ群間比較してしまうと、統計的多重性、つまり α エラーの調整ができないという点がある。 α エラーの調整ができないと、結果を探索的にしか解釈できなくなり、治療効果の差がみられたマーカー・サブグループに対してもう一度検証的な臨床試験を行わなければならないになってしまう。しかしながら、試験前にマーカーを特定していれば、一度の臨床試験で全体比較とマーカー・サブグループ内での比較の両方を検証的に行う試験デザインを立てることができる。もし、試験前にマーカーを特定できない場合でも、'全体で比較→差がない場合にも臨床試験のデータを用いてマーカーを特定→更にその対象に対してサブグループ比較'を一つの臨床試験の中で検証的に行うようにデザインすることも可能である^{8,9)}。

all comers design 実施上の問題として、治療効果が期待できるようなマーカーサブグループ(マーカー(+)とする)が特定されている場合に、そうでないサブグループ(マーカー(−))も含めてランダム化臨床試験を実施するのが困難であると考えられるかもしれない。しかし、マーカー(−)のグループに治療効果がないことが証明さ

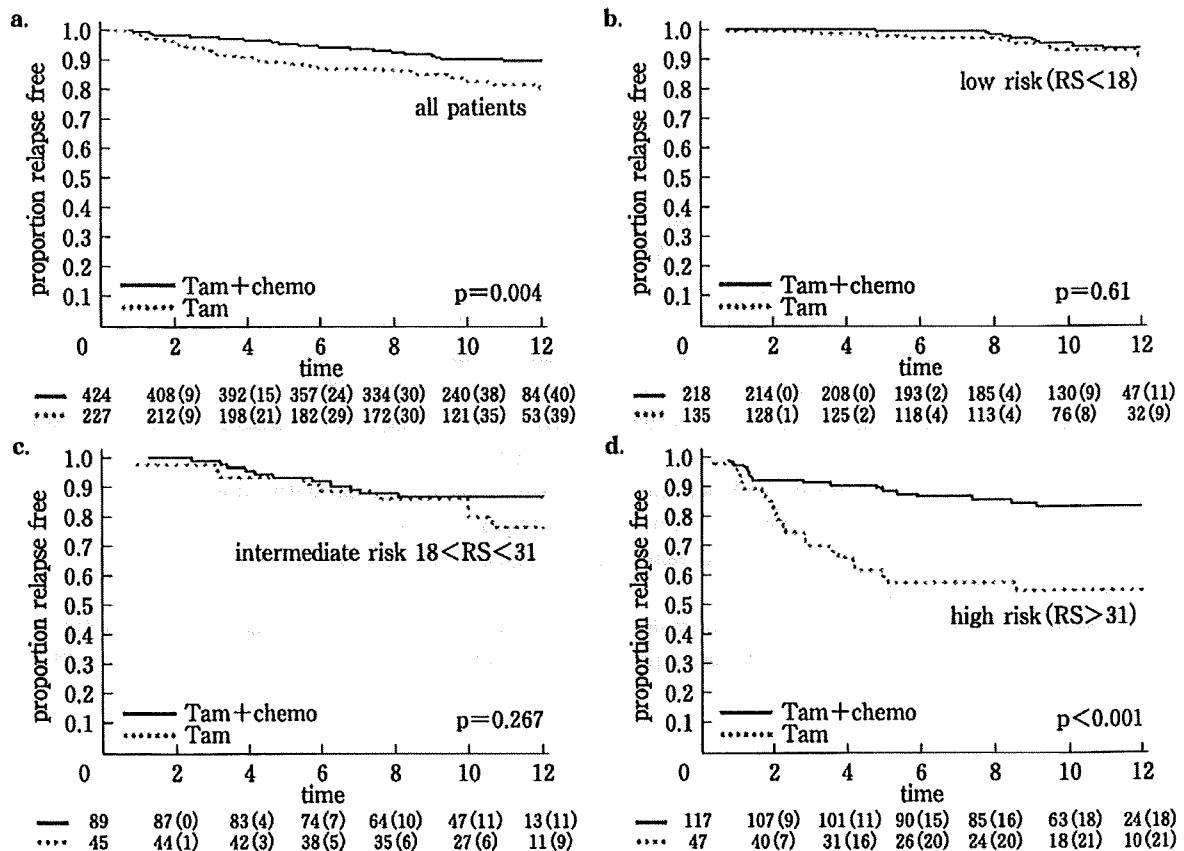


図 6 事後的 all comers design の例

(NSABP B20 対象者の Oncotype DX による再発リスクスコア別治療成績)

数字は at risk の人数, () 内の数字は再発数.

[Reprinted from Paik S, et al: Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 24(23): 3726-3734, 2006. with permission. © 2009 American Society of Clinical Oncology. All rights reserved]

れたわけでなければ、マーカー(-)の対象者を含めたランダム化比較試験を行うことは十分 rationale があると考えられる。

おわりに

分子標的薬をはじめとして、がんの個別化治療開発は全世界的に精力的に取り組まれており、いろいろな考え方、いろいろなデザインで研究が行われている。治療効果を調べたいサブグル

ープを同定し、そのサブグループでの治療効果をランダム化比較試験で検証するというのが基本戦略となると考えられるが、それ以外のサブグループでの治療効果を確認することも重要である。いずれにしろ、治療効果の検証にはランダム化比較試験が必須であるので、予測因子マーカーの探索とその検証を効率よく行うために、質の高い付随研究が実施できるような体制を構築することが重要となると考えられる。

■ 文 献

- 1) Sargent DJ, et al: Clinical trial designs for predictive marker validation in cancer treatment trials. *J Clin Oncol* 23: 2020–2027, 2005.
- 2) Potti A, et al: A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. *N Engl J Med* 355: 570–580, 2006.
- 3) Buyse M, et al: Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst* 98(17): 1183–1192, 2006.
- 4) Kurbacher CM, et al: Use of an ex vivo ATP luminescence assay to direct chemotherapy for recurrent ovarian cancer. *Anticancer Drugs* 9(1): 51–57, 1998.
- 5) Cree IA, et al: A prospective randomized controlled trial of ATP-based tumor chemosensitivity assay (ATP-TCA) directed chemotherapy versus physician's choice in patients with recurrent platinum-resistant ovarian cancer. *Proc ASCO Vol 23, No. 16S, Part I*, 5008, 2005.
- 6) Fisher B, et al: Tamoxifen and chemotherapy for lymph node-negative, estrogen receptor-positive breast cancer. *J Natl Cancer Inst* 89: 1673–1682, 1997.
- 7) Paik S, et al: Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 24(23): 3726–3734, 2006.
- 8) Simon R: Roadmap for developing and validating therapeutically relevant genomic classifiers. *J Clin Oncol* 23: 7332–7341, 2005.
- 9) Freidlin B, Simon R: Adaptive signature design: An adaptive clinical trial design for generating and prospectively testing a gene expression signature for sensitive patients. *Clin Cancer Res* 11: 7872–7878, 2005.

特集

がんの個別化治療を目指したバイオマーカー

バイオマーカーを用いた 臨床試験計画*

山本 精一郎**

Key Words : biomarker, clinical trial, study design

治療開発研究計画における バイオマーカーの役割

バイオマーカーを測定するという事は、それがなんらかの意味で薬剤、もしくは治療効果と関係がある(と期待する)からであろう。分子標的薬であれば、ターゲットで薬が作用していることを測定するためや、マーカーで規定されるサブグループによって治療効果や予後を予測することがバイオマーカーに期待される2つの大きな役割と考えられる。これらは治療開発の段階では異なる役割を担う。順に考えてみたい。

Phase 1の目的は、用量の設定と毒性の評価である。細胞障害性薬剤の場合、毒性と効果が比例しているという性質を利用し、毒性が許容される中で最大の投与量をもっとも効果があることを利用して用量を設定する。つまり、用量制限毒性(DLT)を用いて最大耐用量(MTD)を探索するが、3例コホートがその典型的なデザインである。これに対し、毒性と効果が比例しない可能性のある分子標的薬の場合、phase 1ではそもそも薬剤が作用しているかどうか調べるために、ターゲットと考えられる部位での作用がみられるようなマーカーを測定することが必要となる。しかし、現実にはいいマーカーがなかったり、

作用部位がはっきり特定されなかったりして、マーカーの動きをもって効果を測ることは必ずしも容易ではない。結果として、現実には伝統的な3例コホートデザインを用いることが多いようである。

Phase 2の目的は、有効性によるスクリーニングである。分子標的薬剤では、必ずしも抗腫瘍効果として腫瘍縮小がなくても生存期間の延長が期待できる可能性がある。したがって、腫瘍縮小以外の有効性のマーカーがあればそれを有効性のサロゲートエンドポイントとすることもできるが、サロゲートエンドポイントであるためには有効性の真のエンドポイントである生存期間と相関が高い挙動を示すマーカーである必要があり、そのようなバイオマーカーを薬剤ごと、臓器ごとにphase 2の段階までに開発するのは容易ではない。また、実際には分子標的薬でもある程度の腫瘍縮小効果が期待できるものも多いことから、腫瘍縮小効果をエンドポイントにしたり、生存時間そのものをエンドポイントにしたりすることが一般的である。後者の場合、研究にエントリーした患者背景によって生存期間自体が大きく変わる可能性があるため、標準治療を対象としたランダム化phase 2を行うなど、適切なコントロールの値と比較することが重要である。

Phase 3の目的は真のエンドポイントにおいて

* Clinical trial designs using biomarker.

** Seiichiro YAMAMOTO, Ph.D.: 国立がんセンターがん対策情報センターがん情報・統計部(〒104-0045 東京都中央区築地5-1-1); Cancer Information Services and Surveillance Division, Center for Cancer Control and Information Services, National Cancer Center, Tokyo 104-0045, JAPAN

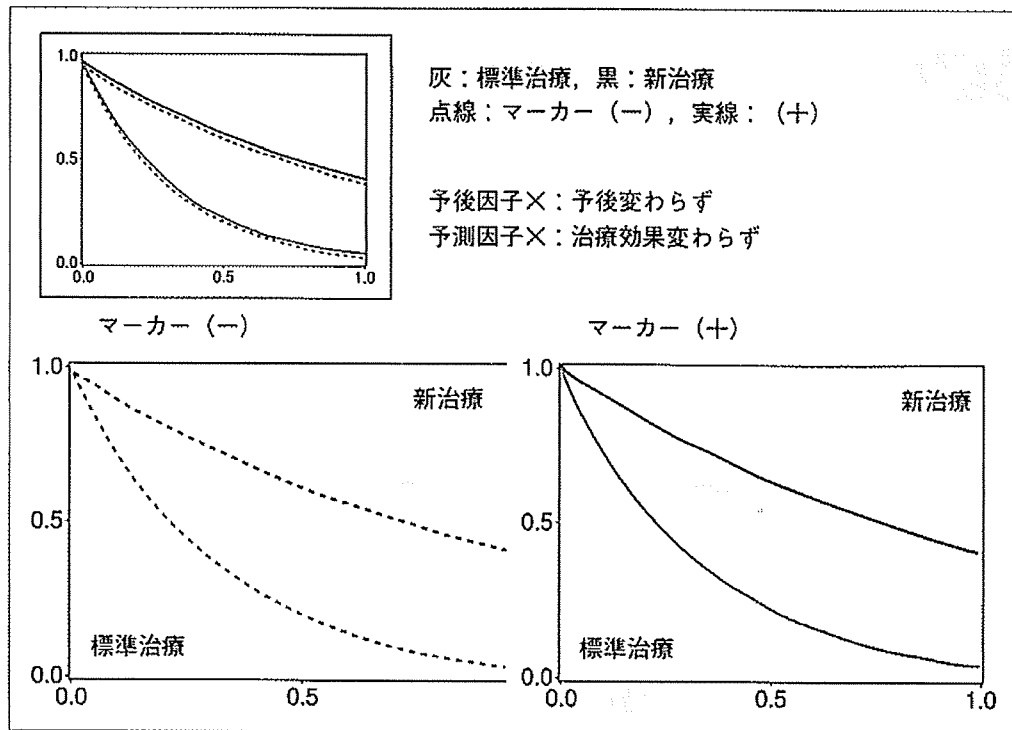


図1 マーカーが予後因子でも予測因子でもない場合

標準治療より優れていることを証明することである。多くの場合、新治療による生存期間が標準治療より長い、もしくは、生存期間においては新治療と標準治療で同等でも、副作用の軽減などその他の大きなメリットが新治療によりもたらされるかどうか、ということになる。すなわち、ターゲットにおいて薬剤が作用しているかという傍証よりも、実際その薬剤で生存期間の延長が認められるか、ということが目的となり、ターゲットでの作用を調べるようなバイオマーカーをエンドポイントとした臨床試験が組まれることは想像しにくい。このコンテキストでのバイオマーカーとは、以下に述べるようなサブグループを特定するようなマーカーのことである。非臨床の研究段階、もしくはphase 1, phase 2での検討で、薬剤がよく作用するようなサブグループを同定することができれば、その集団を対象に治療開発を行うことが効率的と言えるであろう。

予後因子と予測因子を詳しく考える

ここで、あるサブグループをバイオマーカーで特徴づけることを考える。いわゆる、予後因

子マーカーと予測因子マーカーについて詳しく考えてみよう。

予後因子とは、無治療の場合、もしくは標準治療を行った場合に、マーカーで規定されたサブグループ[簡便のためにマーカー(+)とマーカー(-)と呼ぶが、3群以上となってもよい]によって予後が異なるものを言う。しかし、現実には無治療群のデータはあまりないことから標準治療を行った場合の予後を規定する因子のことを言う。これは必然的に標準治療が変わると予後因子も変わる可能性があることを意味している。

予測因子とは、マーカーで規定されたサブグループによって治療効果が異なる因子を言う。たとえば、標準治療ではマーカー(+)と(-)で予後に差がないのに、新治療ではマーカー(+)とマーカー(-)の間で予後に差がある場合、予後の差を生じる要因はマーカーと治療であることから、この因子のことを治療効果の予測因子といい、統計的にはこの状態をマーカーと治療効果に交互作用あり、と言う。これらを生存曲線で表したものが図1~4である。それぞれマーカー(+)(-)のサブグループに、標準治療を行った場合、新治療を行った場合の生存曲線を示し

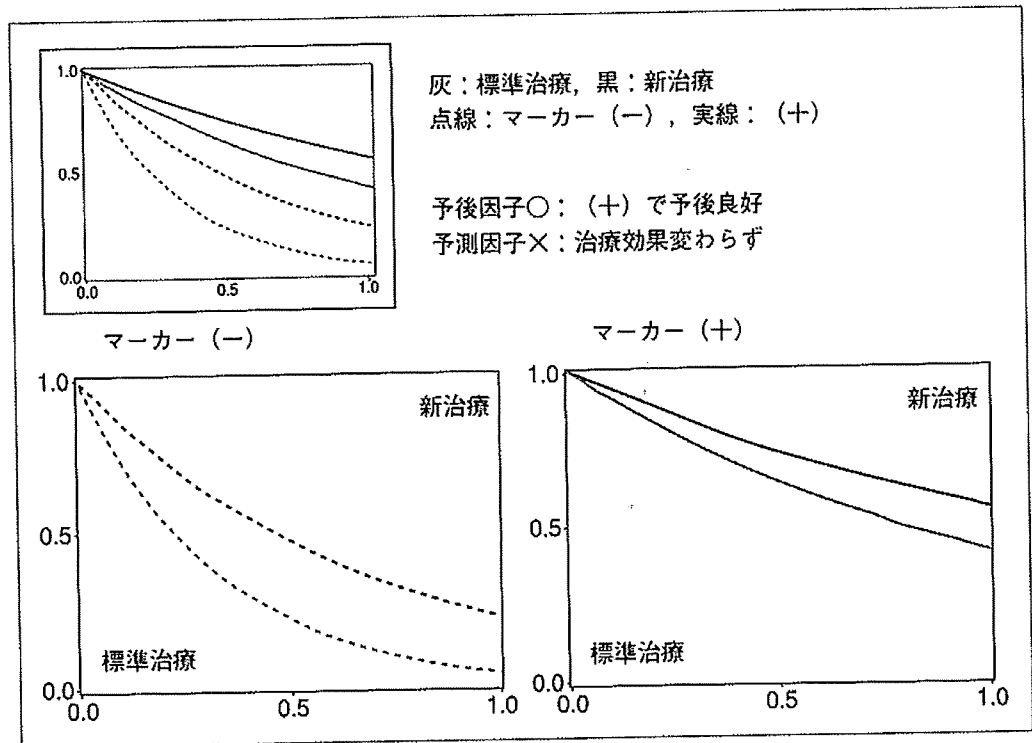


図2 マーカーが予後因子であるが予測因子でない場合

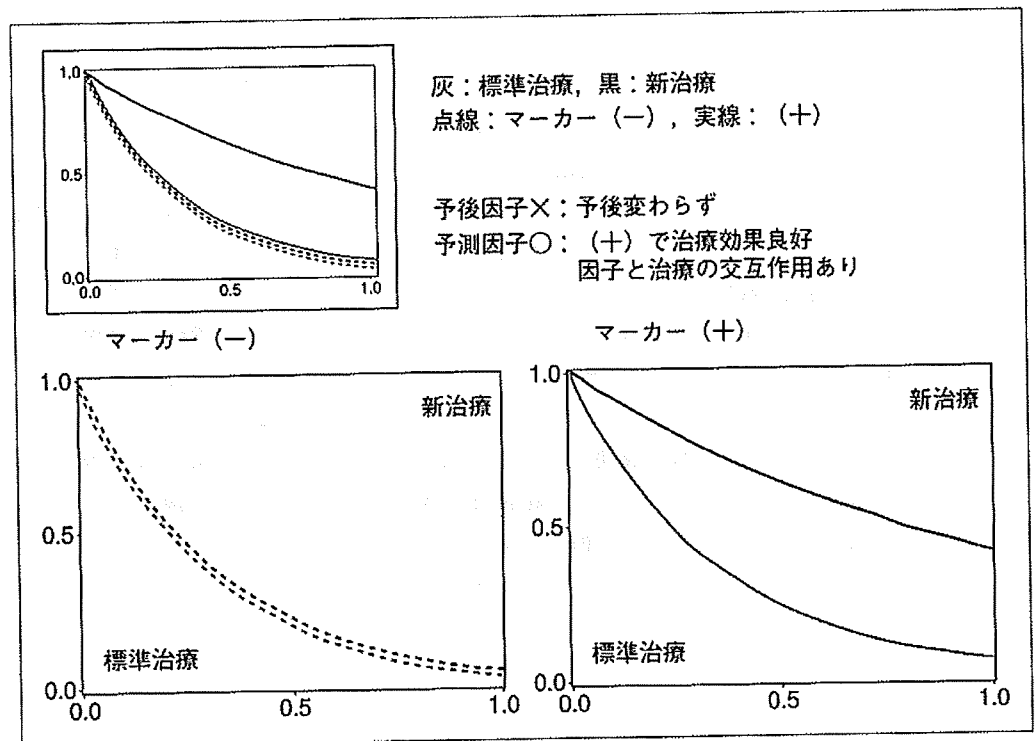


図3 マーカーが予後因子ではないが予測因子である場合

ている。

図1は標準治療群でマーカー(+)とマーカー(-)の間で予後の差がないことから、マーカーは予後因子ではなく、マーカー(+)とマーカー

(-)の間で治療効果に差がないことからマーカーは予測因子でもない。図2は標準治療群でマーカー(+)とマーカー(-)の間で予後に差があるためマーカーは予後因子であるが、マーカー(+)

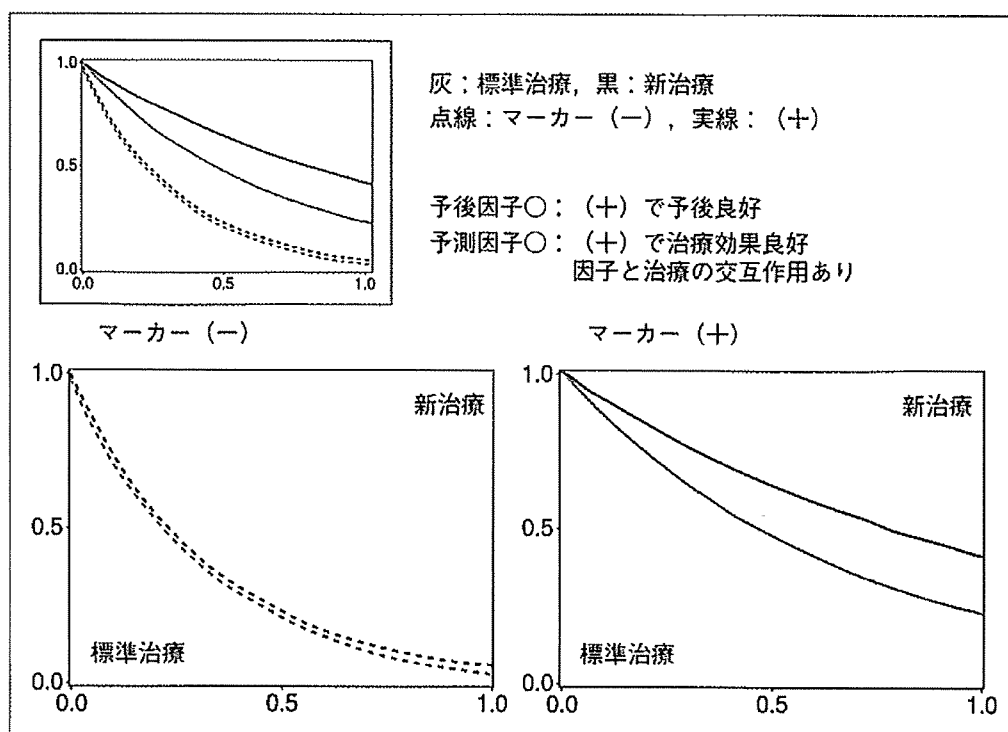


図4 マーカーが予後因子かつ予測因子である場合

とマーカー(-)の間で治療効果に差がないことからマーカーは予測因子ではない。図3は標準治療群でマーカー(+)とマーカー(-)の間で予後に差がないためマーカーは予後因子ではないが、マーカー(+)とマーカー(-)の間で治療効果に差があることからマーカーは予測因子である。図4は標準治療群でマーカー(+)とマーカー(-)の間で予後に差があるためマーカーは予後因子であり、マーカー(+)とマーカー(-)の間で治療効果に差があることからマーカーは予測因子である。これを別の形で表現したものが図5である。縦軸に5年生存率をとり、新治療と無治療で比較している。無治療を標準治療と考えて、図1~4がどれに対応するか考えてほしい。図1はb、図2はe、図3はc、図4はfにそれぞれ対応するのだが、a、bは治療効果がない場合である。この場合でも予後因子は定義できることに注意してほしい。また、マーカーが予測因子である場合で、マーカー(+)で標準治療より新治療で大きな治療効果がある場合でも、マーカー(-)でも新治療の方が少し治療効果がある場合、まったくない場合、逆に標準治療の方が効果がある場合に分けられる。最初の

場合を量的交互作用、あとの2つの場合を質的交互作用と呼ぶ。量的交互作用の場合は、マーカー(-)の群でも治療効果が期待されるため、新治療を導入すべきということになる。質的交互作用ありの場合には、治療効果に差がなければマーカー(-)の群に新治療を導入するかどうかはほかのエンドポイント(副作用など)の成績によるが、標準治療が上回っている場合にはもちろん新治療を導入すべきでないということになる(図6)。

また、治療効果の差で予測因子を定義する場合、あまり指摘されていないが、効果の指標によって予測因子かどうかが変わってしまうという問題がある。たとえば、あるマーカーが予後因子であるとする。マーカー(+)のサブグループにおいて標準治療群での5年生存率が60%だったとして、新治療による5%の治療効果の上乗せがあったとするとこれはハザード比に換算すると0.8となる(生存期間が指数分布に従うことを仮定)。ハザード比は生存期間中央値(MST)の比の逆数と覚えておけばよい。これに対し、マーカー(-)のサブグループにおいて標準治療群での5年生存率が40%の場合、新治療による5%

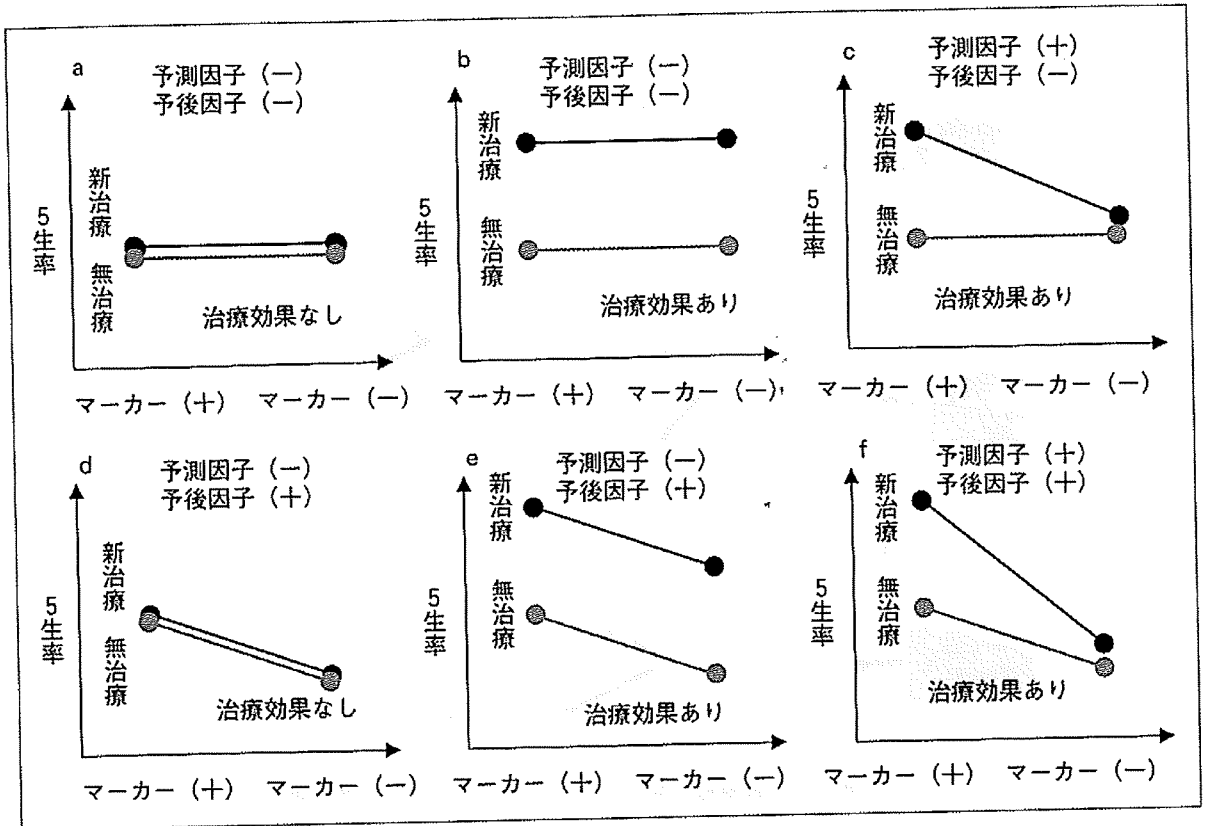


図5 予測因子と予後因子

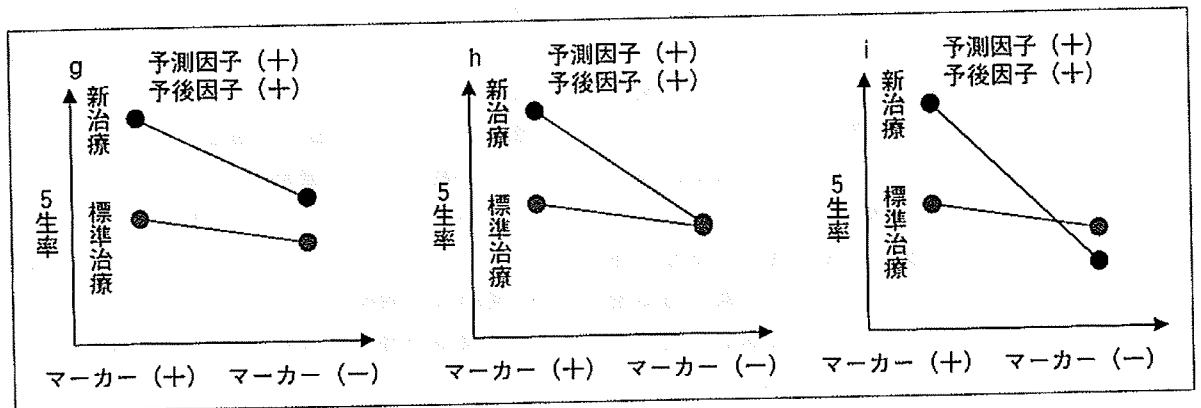


図6 予測因子と治療効果の3つの質的交互作用

の治療効果の上乗せがあったとするとこれはハザード比に換算すると0.7となる。つまり、5年生存率でみるとどちらのサブグループでみても5%の上乗せであるが、ハザード比で考えると治療効果はサブグループ間で異なってしまう。同じ例で、治療効果のハザード比が0.8で一定の場合、マーカー(-)サブグループでの5年生存率の差は6%となる。この場合、治療効果を比で定義するとマーカーは予測因子ではないが、5年生存率の差で定義するとマーカーは予測因子と

いうことになってしまう。どちらが正しいのであろうか。このように統計的交互作用のありなしは用いる効果指標(5年生存率やハザード比など)によって異なってしまうため、疫学では交互作用のことを効果の修飾(effect modification)と呼ぶ。生物学的な交互作用は差の指標における交互作用に対応するという報告もあるが、5年生存率の差は3年生存率の差にも対応しないことを考えると、比の効果指標(通常はハザード比)で交互作用、すなわち予測因子かどうかの判断