ADTree outperformed MLR using an identical dataset. MLR offers some advantages, particularly the use of fewer variables, which facilitates data collection and interpretation of the model. These features of each modelling method represent trade-offs that should be considered when applying the models. The combined use of multiple prediction models could enhance predictive accuracy [27]. We are currently testing the combination of our model and available nomograms in a prospective study.

There are several limitations of this study. Validation using larger databases will more accurately assess the model. The use of many features obtained from imaging studies or physical examination would reduce the number of users depending on the availability of the features. The datasets obtained from multiple institutes would contribute to strict evaluation of the model's versatility whereas such datasets sometimes introduce institute-dependent bias. In this study, we used information from individual pathology reports and the central pathology review is more preferable to evaluate the features in a single criteria. A Web-based interface to facilitate data input and prediction analysis, like the MD Anderson Cancer Centre nomogram, and an automated system to update the model will also be useful. Biomarkers of tumour response, particularly those obtained from midcourse biopsy samples, may increase the predictive accuracy. Integration with subtype-specific biomarkers is also needed to improve the accuracy of the developed model.

In conclusion, we have established a new ADTree-based method to predict pCR after NAC using variables readily collected before NAC. The model could use larger number of variables with keeping high generalization ability and showed the outperformed prediction accuracy compared with MLR as well as was tolerant to missing values and distribution bias in the datasets.

Acknowledgments We thank the doctors and data managers for data collection. We also thank the patients who participated in this study. This study was funded by research grants from the Ministry of Health, Labour and Welfare ("A study on the construction of an algorithm for multimodal therapy with biomarkers for primary breast cancer by formulation of a decision-making process", led by MT, No. H18-3JIGAN-IPPAN-007 and "Reduction and lowering of recurrence risk, toxicity and pharmacoeconomic cost by prediction of efficacy for anticancer agents in breast cancer patients", led by MT; No. H22-GANRINSHO-IPPAN-039), research funds from the Yamagata Prefectural Government and Tsuruoka City, and an International Internship Grant from the Global COE project "Centre for Frontier Medicine", Kyoto University. This study was also supported by the program "Raising Proficient Oncologists" administered by the Japanese Ministry of Education, Culture, Sports, Science and Technology.

Disclosure Dr. Hiroji Iwata has received honoraria from Chugai Pharmaceutical Co., Ltd, Japan. All remaining authors have declared no conflicts of interest.

References

1. van der Hage JH, van de Velde CCJH, Mieog SJSD (2007) Preoperative chemotherapy for women with operable breast cancer. Cochrane Database Syst Rev CD005002. doi:10.1002/14651858.CD005002.pub2
2. Bear HD, Anderson S, Smith RE, Geyer CE Jr, Mamounas EP, Fisher B, Brown AM, Robidoux A, Margolese R, Kahlenberg MS, Paik S, Soran A, Wickerham DL, Wolmark N (2006) Sequential preoperative or postoperative docetaxel added to preoperative doxorubicin plus cyclophosphamide for operable breast cancer: National Surgical Adjuvant Breast and Bowel Project Protocol B-27. J Clin Oncol 24:2019–2027. doi:10.1200/JCO.2005.04.1665
3. Fisher B, Bryant J, Wolmark N, Mamounas E, Brown A, Fisher ER, Wickerham DL, Begovic M, DeCillis A, Robidoux A, Margolese RG, Cruz AB Jr, Hoehn JL, Lees AW, Dimitrov NV, Bear HD (1998) Effect of preoperative chemotherapy on the outcome of women with operable breast cancer. J Clin Oncol 16:2672–2685
4. Mauri D, Pavlidis N, Ioannidis JP (2005) Neoadjuvant versus adjuvant systemic treatment in breast cancer: a meta-analysis. J Natl Cancer Inst 97:188–194. doi:10.1093/jnci/dji021
5. Chang J, Powles TJ, Allred DC, Ashley SE, Clark GM, Makris A, Assersohn L, Gregory RK, Osborne CK, Dowsett M (1999) Biologic markers as predictors of clinical outcome from systemic therapy for primary operable breast cancer. J Clin Oncol 17:3058–3063
6. Colleoni M, Viale G, Goldhirsch A (2009) Lessons on responsiveness to adjuvant systemic therapies learned from the neoadjuvant setting. Breast 18(Suppl 3):S137–S140. doi:10.1016/S0960-9776(09)70289-9
7. von Minckwitz G, Untch M, Nuesch E, Loibl S, Kaufmann M, Kummel S, Fasching PA, Eiermann W, Blohmer JU, Costa SD, Mehta K, Hilfrich J, Jackisch C, Gerber B, du Bois A, Huober J, Hanusch C, Konecny G, Fett W, Stickeler E, Harbeck N, Muller V, Juni P (2011) Impact of treatment characteristics on response of different breast cancer phenotypes: pooled analysis of the German neo-adjuvant chemotherapy trials. Breast Cancer Res Treat 125:145–156. doi:10.1007/s10549-010-1228-x
8. Colleoni M, Viale G, Zahrieh D, Bottiglieri L, Gelber RD, Veronesi P, Balduzzi A, Torrisi R, Luini A, Intra M, Dellapasqua S, Cardillo A, Ghisini R, Peruzzotti G, Goldhirsch A (2008) Expression of ER, PgR, HER1, HER2, and response: a study of preoperative chemotherapy. Ann Oncol 19:465–472. doi:10.1093/annonc/mdm509
9. Darb-Esfahani S, Loibl S, Muller BM, Roller M, Denkert C, Komor M, Schluns K, Blohmer JU, Budczies J, Gerber B, Noske A, du Bois A, Weichert W, Jackisch C, Dietel M, Richter K, Kaufmann M, von Minckwitz G (2009) Identification of biology-based breast cancer types with distinct predictive and prognostic features: role of steroid hormone and HER2 receptor expression in patients treated with neoadjuvant anthracycline/taxane-based chemotherapy. Breast Cancer Res 11:R69. doi:10.1186/bcr2363
10. Jones RL, Salter J, A'Hern R, Nerurkar A, Parton M, Reis-Filho JS, Smith IE, Dowsett M (2010) Relationship between oestrogen receptor status and proliferation in predicting response and long-term outcome to neoadjuvant chemotherapy for breast cancer. Breast Cancer Res Treat 119:315–323. doi:10.1007/s10549-009-0329-x
11. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ (2011) Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast

Cancer 2011. Ann Oncol 22:1736–1747. doi:10.1093/annonc/mdr304
12. Colleoni M, Bagnardi V, Rotmensz N, Viale G, Mastropasqua M, Veronesi P, Cardillo A, Torrisi R, Luini A, Goldhirsch A (2010) A nomogram based on the expression of Ki-67, steroid hormone receptors status and number of chemotherapy courses to predict pathological complete remission after preoperative chemotherapy for breast cancer. Eur J Cancer 46:2216–2224. doi:10.1016/j.ejca.2010.04.008
13. Rouzier R, Pusztai L, Delaloge S, Gonzalez-Angulo AM, Andre F, Hess KR, Buzdar AU, Garbay JR, Spielmann M, Mathieu MC, Symmans WF, Wagner P, Atallah D, Valero V, Berry DA, Hortobagyi GN (2005) Nomograms to predict pathologic complete response and metastasis-free survival after preoperative chemotherapy for breast cancer. J Clin Oncol 23:8331–8339. doi:10.1200/JCO.2005.01.2898
14. Lee JK, Coutant C, Kim YC, Qi Y, Theodorescu D, Symmans WF, Baggerly K, Rouzier R, Pusztai L (2010) Prospective comparison of clinical and genomic multivariate predictors of response to neoadjuvant chemotherapy in breast cancer. Clin Cancer Res 16:711–718. doi:10.1158/1078-0432.CCR-09-2247
15. Tabchy A, Valero V, Vidaurre T, Lluch A, Gomez H, Martin M, Qi Y, Barajas-Figueroa LJ, Souchon E, Coutant C, Doimi FD, Ibrahim NK, Gong Y, Hortobagyi GN, Hess KR, Symmans WF, Pusztai L (2010) Evaluation of a 30-gene paclitaxel, fluorouracil, doxorubicin, and cyclophosphamide chemotherapy response predictor in a multicenter randomized trial in breast cancer. Clin Cancer Res 16:5351–5361. doi:10.1158/1078-0432.CCR-10-1265
16. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, Ravdin P, Bugarini R, Baehner FL, Davidson NE, Sledge GW, Winer EP, Hudis C, Ingle JN, Perez EA, Pritchard KI, Shepherd L, Gralow JR, Yoshizawa C, Allred DC, Osborne CK, Hayes DF (2010) Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol 11:55–65. doi:10.1016/S1470-2045(09)70314-6
17. Liedtke C, Hatzis C, Symmans WF, Desmedt C, Haibe-Kains B, Valero V, Kuerer H, Hortobagyi GN, Piccart-Gebhart M, Sotiriou C, Pusztai L (2009) Genomic grade index is associated with response to chemotherapy in patients with breast cancer. J Clin Oncol 27:3185–3191. doi:10.1200/JCO.2008.18.5934
18. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, Palazzo J, Marron JS, Nobel AB, Mardis E, Nielsen TO, Ellis MJ, Perou CM, Bernard PS (2009) Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 27:1160–1167. doi:10.1200/JCO.2008.18.1370
19. Straver ME, Glas AM, Hannemann J, Wesseling J, van de Vijver MJ, Rutgers EJ, Vrancken Peeters MJ, van Tinteren H, Van't Veer LJ, Rodenhuis S (2010) The 70-gene signature as a response predictor for neoadjuvant chemotherapy in breast cancer. Breast Cancer Res Treat 119:551–558. doi:10.1007/s10549-009-0333-1
20. Freund Y, Mason L (1999) The alternating decision tree learning algorithm. In: Proceedings of the sixteenth international conference on machine learning, pp 124–133
21. Che D, Hockenbury C, Marmelstein R, Rasheed K (2010) Classification of genomic islands using decision trees and their ensemble algorithms. BMC Genomics 11(Suppl 2):S1. doi:10.1186/1471-2164-11-S2-S1
22. von Minckwitz G, Rezai M, Loibl S, Fasching PA, Huober J, Tesch H, Bauerfeind I, Hilfrich J, Eidtmann H, Gerber B, Hanusch C, Kühn T, du Bois A, Blohmer JU, Thomssen C, Dan Costa S, Jackisch C, Kaufmann M, Mehta K, Untch M (2010) Capecitabine in addition to anthracycline- and taxane-based neoadjuvant treatment in patients with primary breast cancer: phase III GeparQuattro study. J Clin Oncol 28:2015–2023. doi:10.1200/JCO.2009.23.8303
23. Jinno H, Sakata M, Hayashida T, Takahashi M, Sato T, Seki H, Kitagawa Y (2011) Primary systemic chemotherapy of breast cancer: indication and predictive factors. Breast Cancer 18:74–79. doi:10.1007/s12282-010-0226-5
24. von Minckwitz G, Blohmer JU, Raab G, Löhr A, Gerber B, Heinrich G, Eidtmann H, Kaufmann M, Hilfrich J, Jackisch C, Zuna I, Costa SD (2005) In vivo chemosensitivity-adapted preoperative chemotherapy in patients with early-stage breast cancer: the GEPARTRIO pilot study. Ann Oncol 16:56–63. doi:10.1093/annonc/mdi001
25. Perou CM (2010) Molecular stratification of triple-negative breast cancers. Oncologist 15(Suppl 5):39–48. doi:10.1634/theoncologist.2010-S5-39
26. Rouzier R, Pusztai L, Garbay JR, Delaloge S, Hunt KK, Hortobagyi GN, Berry D, Kuerer HM (2006) Development and validation of nomograms for predicting residual tumor size and the probability of successful conservative surgery with neoadjuvant chemotherapy for breast cancer. Cancer 107:1459–1466. doi:10.1002/cncr.22177
27. Fan C, Prat A, Parker JS, Liu Y, Carey LA, Troester MA, Perou CM (2011) Building prognostic models for breast cancer patients using clinical variables and hundreds of gene expression signatures. BMC Med Genomics 4:3. doi:10.1186/1755-8794-4-3

Takada et al. BMC Medical Informatics and Decision Making 2012, 12:54
http://www.biomedcentral.com/1472-6947/12/54

BMC Medical Informatics & Decision Making

RESEARCH ARTICLE Open Access

Prediction of axillary lymph node metastasis in primary breast cancer patients using a decision tree-based model

Masahiro Takada[1,2], Masahiro Sugimoto[3,4,5], Yasuhiro Naito[3,4,6], Hyeong-Gon Moon[7], Wonshik Han[7], Dong-Young Noh[7], Masahide Kondo[8], Katsumasa Kuroi[9], Hironobu Sasano[10], Takashi Inamoto[11], Masaru Tomita[3,4,6] and Masakazu Toi[1*]

Abstract

Background: The aim of this study was to develop a new data-mining model to predict axillary lymph node (AxLN) metastasis in primary breast cancer. To achieve this, we used a decision tree-based prediction method—the alternating decision tree (ADTree).

Methods: Clinical datasets for primary breast cancer patients who underwent sentinel lymph node biopsy or AxLN dissection without prior treatment were collected from three institutes (institute A, $n = 148$; institute B, $n = 143$; institute C, $n = 174$) and were used for variable selection, model training and external validation, respectively. The models were evaluated using area under the receiver operating characteristics (ROC) curve analysis to discriminate node-positive patients from node-negative patients.

Results: The ADTree model selected 15 of 24 clinicopathological variables in the variable selection dataset. The resulting area under the ROC curve values were 0.770 [95% confidence interval (CI), 0.689–0.850] for the model training dataset and 0.772 (95% CI: 0.689–0.856) for the validation dataset, demonstrating high accuracy and generalization ability of the model. The bootstrap value of the validation dataset was 0.768 (95% CI: 0.763–0.774).

Conclusions: Our prediction model showed high accuracy for predicting nodal metastasis in patients with breast cancer using commonly recorded clinical variables. Therefore, our model might help oncologists in the decision-making process for primary breast cancer patients before starting treatment.

Keywords: Breast cancer, Lymph node metastasis, Data mining, Alternating decision tree

Background

Axillary lymph node (AxLN) metastasis is one of the most important prognostic factors in patients with primary breast cancer for predicting survival [1-4]. Sentinel lymph node (SLN) biopsy is widely used to determine AxLN status and avoids AxLN dissection (ALND). However, SLN biopsy is an invasive procedure. Therefore, predicting AxLN metastasis before SLN biopsy using commonly recorded clinical variables would be helpful for oncologists and could avoid this procedure, especially in elderly patients or patients with complications. Consequently, many mathematical models have been developed to predict AxLN metastasis, including nomograms and scoring systems [5-14]. For example, the Memorial Sloan-Kettering Cancer Center (MSKCC) developed a nomogram to predict the presence of SLN metastasis [6] that is now used worldwide.

Technically, nomograms use multiple logistic regression (MLR) to predict a binary outcome based on a combination of risk factors. This well-established method has a limitation in that it incorporates only a few independent variables so that the model can accurately predict risk in independent datasets, by avoiding over-fitting to the given datasets. Such prediction models should also tolerate missing values, which are common in clinical

* Correspondence: toi@kuhp.kyoto-u.ac.jp
[1]Department of Breast Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan
Full list of author information is available at the end of the article

datasets. Thus, new methods to cope with a greater number of variables and that provide accurate prediction and robustness against missing values are required.

Machine learning has been applied to problems across many fields, including bioinformatics [15], and it is thought to overcome or reduce the impact of the limitations of MLR. Here, we used the alternating decision tree (ADTree) [16,17] as a core algorithm. This algorithm consists of a root node and multiple simple decision trees in which an index is associated with each leaf node, and its final predictive value is the sum of the indices of the leaf nodes fulfilling the patients' condition. This algorithm also differs from standard 'if–then' decision trees and classification and regression trees (CART). The ADTree method has several advantages compared with these other machine learning algorithms, including: (1) several comparative studies have shown higher accuracy and versatility for ADTree than other machine learning methods [18,19]; and (2) the ADTree model structure is less complex than other methods [16], which facilitates model interpretation and reduces the need for model optimization.

The purpose of this study was to develop a new mathematical model to predict AxLN metastasis in patients with primary breast cancer using preoperative clinicopathological information.

Methods

Patients

The training datasets consisted of consecutive patients who were treated at two institutions in Japan. Patients with histologically confirmed primary invasive breast cancer who underwent SLN biopsy or ALND without prior treatment were eligible for this study. We included patients whose maximum tumor size was ≤ 4 cm. We identified 148 patients from the Tokyo Metropolitan Cancer and Infectious Diseases Centre Komagome Hospital who were treated between 2005 and 2006 (Tokyo dataset) and 143 patients from Kyoto University Hospital treated between 2008 and 2009 (Kyoto dataset).

The external validation dataset was collected from Seoul National University Hospital, Korea, and consisted of patients consecutively treated between January 6, 2010, and April 16, 2010 (Seoul dataset). We included 174 patients who underwent SLN biopsy and met the same eligibility criteria as the modeling dataset. All datasets were collected after establishing the methodology for SLN biopsy, and no significant difference in SLN biopsy accuracies was expected [20,21].

The study protocol was approved by the institutional review board at Kyoto University Hospital. All patient data were anonymized and allocated numbers according to Japanese ethical guidelines for epidemiologic research.

Data collection and sentinel lymph node biopsy

Clinical data collected included age, body mass index (BMI), menopausal status, physical findings (based on inspection or palpation), diagnostic mammography and ultrasonography findings, pathological findings from needle biopsy before treatment (*e.g.*, histological type, histological/nuclear grade, estrogen receptor status, progesterone receptor status, and human epidermal growth factor receptor 2 [HER2] status), and type of axillary surgical procedure (SLN biopsy or ALND) as predictive variables. Pathological findings from surgical specimens (presence or absence of lymph node metastasis) were used as outcome variables for prediction by the ADTree model. All data were retrospectively collected from databases maintained at each institution.

The grading criteria were established by a committee of specialists from the fields of breast surgery, diagnostic radiology and pathology. We reviewed all of the images from which mammographic and ultrasonographic variables were obtained, and these parameters were determined using Japanese diagnostic guidelines for mammography and ultrasonography based on the American College of Radiology Breast Imaging Reporting and Data System [22]. These variables were reviewed by physicians certified for imaging diagnosis by the relevant accreditation organizations in Japan.

The techniques used for SLN biopsy and histological evaluations are described elsewhere [21]. In the Tokyo dataset, SLNs were identified using a radioactive tracer (^{99m}Tc-phytate). In the Kyoto dataset, they were identified using blue dye and a fluorescence navigation technique using indocyanine green. In the Seoul dataset, SLNs were identified using both blue dye and a radioactive tracer. At each institution, the SLNs were step-sectioned, stained with hematoxylin and eosin (H&E), and diagnosed by trained pathologists. Lymph nodes obtained after ALND were evaluated using a single H&E-stained section from each node. Metastases were defined as the presence of a tumor deposit > 0.2 mm in diameter in at least one lymph node. Several clinical trials have reported no significant differences in the identification rate or accuracy of SLN methodologies [20,23,24].

Data analysis

A summary of the model development and validation procedure is shown in Appendix A (Additional file 1). The model development phase consisted of three steps. First, bias-control virtual datasets were generated from the Tokyo dataset by randomly selecting individuals allowing for redundant selection. These datasets contained an approximately equal ratio of patients negative and positive for AxLN. Second, a prediction model containing multiple ADTrees was trained on a generated dataset, and the mean value of the individual trees' predictions values was

used to enhance the accuracy and generalization ability in a process referred to as the ensemble technique [25]. This model development procedure was repeated for different modeling conditions, *e.g.* the number of nodes, and all virtual datasets. Third, we selected the model yielding the best area under the receiver operating characteristics (ROC) curves (AUC) value with the Kyoto dataset. Finally, we performed external validation of the chosen model using the Seoul dataset.

The established model was further evaluated as follows. First, we performed bootstrap analysis using the Seoul dataset to obtain unbiased estimates of the developed model. Second, the relative importance of the variables in the model was analyzed by randomly changing the values of each variable (sensitivity analysis). Third, missing values in the Seoul datasets were changed to random values to evaluate the model's tolerance against missing values (missing value analysis). Fourth, the number of trees in the prediction model was reduced to evaluate the relationship between the number of variables in the model and the prediction accuracy (pruning analysis).

Two hundred bias-controlled datasets were generated using different random values. The number of nodes (called boosting iterations) in an ADTree was expanded from 10, 11, ... to 20 in each trial. For the ensemble procedure, we randomly sampled individuals to generate multiple datasets, and the averaged prediction of the trained models for each dataset was used [26]. In this ensemble procedure, the number of ADTrees ranged from 2, 3, ..., to 20, with a random seed to generate random values (1, 2, ..., and 10). Two hundred replicates with different random values were generated for each bootstrap, sensitivity and missing value analysis.

Weka (ver. 3.6.1; University of Waikato, Hamilton, NZ) [27] was used for resampling, the ensemble procedure and ADTree development. The Mann–Whitney test and AUCs with 95% confidence interval (CI) were calculated using GraphPad Prism version 5.04 (GraphPad Software, Inc., San Diego, CA). JMP® (ver. 7.0.1, SAS Institute, Cary, NC, USA) was used for other statistical analyses.

Results

The clinicopathological characteristics of patients in each dataset are summarized in Table 1. The proportion of patients with AxLN metastasis was 29.7%, 30.8% and 23.6% in the Tokyo, Kyoto and Seoul datasets. The proportion of patients with AxLN metastasis in the Seoul dataset was not significantly different from the other datasets ($P = 0.292$).

The model with the best AUC value in the Kyoto dataset included five ADTrees with 13 nodes (Figure 1 and Appendix B (Additional file 1)). A total of 15 variables were included: age, BMI, seven ultrasonographic variables (maximum tumor size, tumor depth/width ratio, multifocality, echogenic halo, interruption of the anterior border of the mammary gland, maximum size of lymph nodes, and a loss of hilum in lymph nodes), two mammographic variables (shape and distribution of calcification), two physical examination variables (skin dimpling and nipple discharge) and two pathological variables (histological/nuclear grade, HER2 status). The method used to calculate the score is shown in Appendix C (Additional file 1).

The ROC curves for each dataset are shown in Figure 2. The AUC values were 0.917 (95% CI: 0.871–0.964, $P < 0.0001$) for the Tokyo dataset, 0.770 (95% CI: 0.689–0.850, $P < 0.0001$) for the Kyoto dataset and 0.772 (95% CI: 0.689–0.856, $P < 0.0001$) for the Seoul dataset. Box plots of the predicted probabilities of AxLN metastasis are shown in Figure 3. The model discriminated node-positive patients from node-negative patients at statistically significant levels ($P < 0.0001$), although there was some overlap of the predicted probability distribution of node-negative and node-positive status in each dataset.

The mean AUC values yielded by bootstrap analysis remained high for each dataset, being 0.916 (95% CI: 0.913–0.919), 0.766 (95% CI: 0.760–0.772) and 0.768 (95% CI: 0.763–0.774) for the Tokyo, Kyoto and Seoul datasets, respectively. A calibration plot of the model developed using the Kyoto and Seoul datasets is shown in Appendix D (Additional file 1). The predicted probabilities were divided into quintiles according to their values, and the mean and actual frequencies of AxLN metastasis were plotted for each quintile.

In the sensitivity analysis, the AUC values decreased remarkably when the following variables were randomly replaced: echogenic halo, maximum size of the lymph nodes, maximum size of the tumor, skin dimpling, and interruption of the anterior border of the mammy gland. This indicates that the developed model was more sensitive to this variable than the other variables, which hardly affected AUC values (Figure 4). In the missing value analysis, 33 and 19 patients with missing values were selected from the Kyoto and Seoul datasets, and we validated the developed model by replacing missing values with random values. This procedure was repeated 200 times for each dataset, and the mean AUC values were 0.884 (95% CI: 0.882–0.887) and 0.688 (95% CI: 0.684–0.692) for the Kyoto and Seoul datasets, respectively. In the pruning analysis, the number of trees was reduced from 5 to 1, and AUC values were calculated for the Tokyo datasets in cross-validation mode, in addition to the Kyoto and Seoul datasets (Appendix E (Additional file 1)).

The predictive performance of the MSKCC nomogram and a scoring system developed at Russells Hall Hospital, United Kingdom, were evaluated using the Seoul dataset

Table 1 Patient characteristics and incidence of lymph node metastasis

Variables	Tokyo dataset		Kyoto dataset		Seoul dataset		*P*-value[§]
	No	%	No	%	No	%	
No. of patients	148	(100)	143	(100)	174	(100)	
Age							<0.001
Median	55		60		50		
Range	(31–85)		(26–88)		(25–74)		
Body mass index							0.019
Median	22.9		22.3		23.2		
Range	(16.6–43.2)		(14.8–31.4)		(17.8–37)		
Unknown	3	(2)	0	(0)	1	(0.6)	
Clinical T classification							0.2621
T1	102	(68.9)	100	(69.9)	108	(62.1)	
T2	46	(31.1)	43	(30.1)	66	(37.9)	
Clinical N classification							0.002
N0	137	(92.6)	135	(94.4)	174	(100)	
N1	11	(7.4)	8	(5.6)	0	(0)	
Skin dimpling							<0.001
Yes	22	(14.9)	14	(9.8)	2	(1.1)	
No	109	(73.6)	129	(90.2)	172	(98.9)	
Unknown	17	(11.5)	0	(0)	0	(0)	
Nipple discharge							0.238
Yes	6	(4.1)	2	(1.4)	3	(1.7)	
No	138	(93.2)	141	(98.6)	170	(97.7)	
Unknown	4	(2.7)	0	(0)	1	(0.6)	
Mammography							
Presence of masses							0.284
Yes	90	(60.8)	88	(61.5)	102	(58.6)	
Focal asymmetry	22	(14.9)	20	(14)	39	(22.4)	
No	35	(23.6)	26	(18.2)	33	(19)	
Unknown	1	(0.7)	9	(6.3)	0	(0)	
Presence of calcifications							0.037
Yes	67	(45.3)	44	(30.8)	59	(33.9)	
No	81	(54.7)	94	(65.7)	115	(66.1)	
Unknown	0	(0)	5	(3.5)	0	(0)	
Shape of calcifications							0.010
Fine branching or casting	4	(6)	1	(2.3)	3	(5.1)	
Pleomorphic	9	(13.4)	11	(25)	21	(35.6)	
Amorphous or indistinct	43	(64.2)	27	(61.4)	35	(59.3)	
Round or benign	11	(16.4)	4	(9.1)	0	(0)	
Unknown	0	(0)	1	(2.3)	0	(0)	
Distribution of calcifications							0.024
Linear or segmented	26	(38.8)	14	(31.8)	22	(37.3)	
Grouped or clustered	30	(44.8)	29	(65.9)	36	(61)	
Regional or diffuse	9	(13.4)	1	(2.3)	1	(1.7)	
Unknown	2	(3)	0	(0)	0	(0)	

Table 1 Patient characteristics and incidence of lymph node metastasis *(Continued)*

Ultrasonography							
Presence of masses							0.264
Yes	142	(95.9)	133	(93)	161	(92.5)	
No	5	(3.4)	10	(7)	13	(7.5)	
Unknown	1	(0.7)	0	(0)	0	(0)	
Multifocality							0.114
Yes	27	(19)	14	(10.5)	21	(13)	
No	115	(81)	119	(89.5)	140	(87)	
Maximum tumor size (mm)							0.004
Median	16		16.1		19		
Range	(4–37)		(5–35)		(4–37)		
Depth/width ratio							0.001
Median	0.72		0.67		0.64		
Range	(0.31–1.36)		(0.22–1.43)		(0.33–1.27)		
Unknown	0	(0)	9	(6.8)	0	(0)	
Echogenic halo							<0.001
Yes	32	(22.5)	62	(46.6)	38	(23.6)	
No	109	(76.8)	71	(53.4)	123	(76.4)	
Unknown	1	(0.7)	0	(0)	0	(0)	
Interruption of the anterior border of the mammary gland							0.807
Yes	99	(69.7)	91	(68.4)	106	(65.8)	
No	43	(30.3)	42	(31.6)	54	(33.5)	
Unknown	0	(0)	0	(0)	1	(0.6)	
Detection of LNs							0.130
Detectable	49	(33.1)	37	(25.9)	56	(32.2)	
Not detectable	82	(55.4)	105	(73.4)	117	(67.2)	
Unknown	17	(11.5)	1	(0.7)	1	(0.6)	
Maximum size (mm) of LNs							0.010
Median	11		10		10		
Range	(5–22)		(3–32)		(4–17)		
Unknown	0	(0)	4	(10.8)	1	(1.8)	
Hilum of LNs							0.021
Detectable	43	(87.8)	27	(73)	36	(64.3)	
Not detectable	6	(12.2)	9	(24.3)	20	(35.7)	
Unknown	0	(0)	1	(2.7)	0	(0)	
Histological type							0.584
Invasive ductal carcinoma	135	(91.2)	129	(90.2)	160	(92)	
Invasive lobular carcinoma	5	(3.4)	3	(2.1)	7	(4)	
Other specific types	8	(5.4)	11	(7.7)	7	(4)	
Estrogen receptor[†]							0.023
Positive	119	(80.4)	114	(79.7)	121	(69.5)	
Negative	27	(18.2)	29	(20.3)	53	(30.5)	
Unknown	2	(1.4)	0	(0)	0	(0)	
Progesterone receptor[†]							0.427
Positive	83	(56.1)	89	(62.2)	96	(55.2)	

Table 1 Patient characteristics and incidence of lymph node metastasis *(Continued)*

Negative	63	(42.6)	54	(37.8)	78	(44.8)	
Unknown	2	(1.4)	0	(0)	0	(0)	
HER2[‡]							0.019
Positive	18	(12.2)	11	(7.7)	29	(16.7)	
Negative	121	(81.8)	131	(91.6)	125	(71.8)	
Unknown	9	(6.1)	1	(0.7)	20	(11.5)	
Histological/nuclear grade							<0.001
1	64	(43.2)	43	(30.1)	4	(2.3)	
2	47	(31.8)	63	(44.1)	82	(47.1)	
3	27	(18.2)	36	(25.2)	88	(50.6)	
Unknown	10	(6.8)	1	(0.7)	0	(0)	
LN metastasis							0.292
Yes	44	(29.7)	44	(30.8)	41	(23.6)	
No	104	(70.3)	99	(69.2)	133	(76.4)	

Note:
Abbreviations: LN, lymph node.
[†]Estrogen receptor or progesterone receptor positive was defined as ≥10% positively stained cells on immunohistochemical (IHC) testing.
[‡]HER2 positive was defined as IHC 3+ or positive on fluorescence *in situ* hybridization testing.
[§]The χ^2 test or Kruskal–Wallis test was used depending on the distribution of patients in each variable and dataset.

[6,28]. Both models included lymphovascular invasion (LVI) as an input variable. However, LVI is not routinely reported for needle biopsy samples because of its uncertain diagnostic role [29]. As preoperative pathological diagnosis in the Seoul dataset was performed by needle biopsy, we used LVI status assessed on surgical specimens. The resulting AUC values were 0.664 (95% CI; 0.560–0.768, $P = 0.0033$) for the nomogram and 0.620 (95% CI; 0.509–0.731, $P = 0.0032$) for the scoring system using individuals without missing values ($n = 131$)

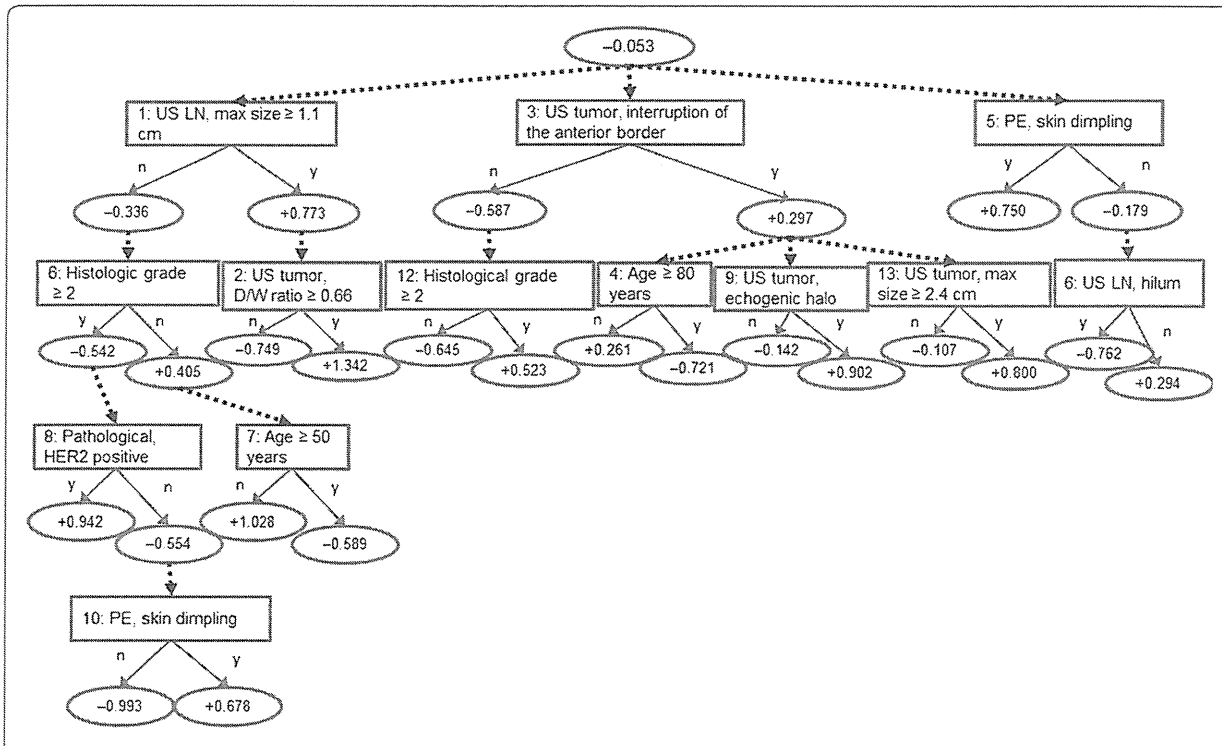


Figure 1 ADTree model. The final prediction model consisted of five ADTree-based prediction models; the other four models are depicted in Appendix B (Additional file 1). The method used to calculate the prediction score for each model is shown in Appendix C (Additional file 1). The final prediction was calculated by calculating the mean score of the five ADTree models.

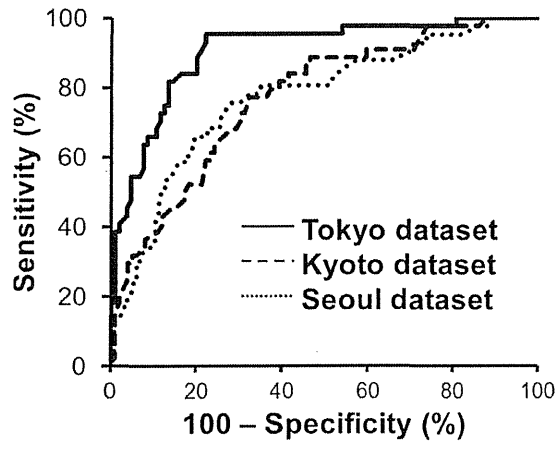


Figure 2 Receiver operating characteristic (ROC) curves of the prediction model. The area under the ROC curve (AUC) values were 0.917 (95% CI: 0.871–0.964, $P < 0.0001$), 0.770 (95% CI: 0.689–0.850, $P < 0.0001$) and 0.772 (95% CI: 0.689–0.856, $P < 0.0001$) for the Tokyo, Kyoto and Seoul (validation dataset) datasets, respectively.

(Appendix F (Additional file 1)). The AUC value using the corresponding patients in the Seoul dataset was 0.777 (95% CI: 0.689–0.864, $P < 0.001$) for ADTree.

Discussion

A data-mining model generated using the ADTree ensemble technique improved the prediction of AxLN metastasis in patients with primary breast cancer, compared with older models such as the MSKCC nomogram. Evaluation using an external validation dataset and bootstrap analysis revealed high AUC values of 0.772 and

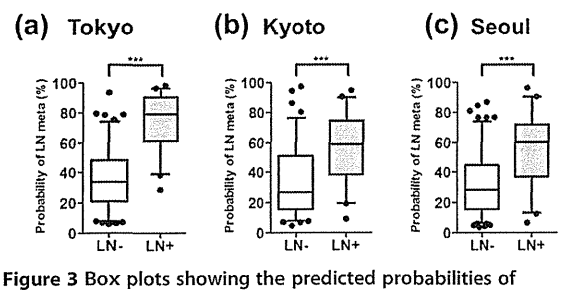


Figure 3 Box plots showing the predicted probabilities of lymph node metastasis for the Tokyo (a), Kyoto (b) and Seoul (c) datasets. In each figure, the boxes show the actual number of lymph node-negative (LN–) and -positive (LN+) patients, respectively. The whisker box-plots indicate the 5th, 25th, 50th, 75th and 95th percentiles (from the bottom bar to the upper bar) of the predicted probabilities. The probabilities <5% and >95% are plotted individually. The differences between LN– and LN + were statistically significant ($P < 0.0001$; Mann–Whitney test) in all datasets. The median predicted probabilities of LN– and LN + were **(a)** 33.5 (95% CI: 31.8–39.4) and 78.9 (95% CI: 69.3–80.4), **(b)** 33.6 (95% CI: 29.1–38.0) and 58.9 (95% CI: 49.3–62.9), and **(c)** 32.3 (95% CI: 28.8–35.8) and 59.9 (95% CI: 48.2–62.6).

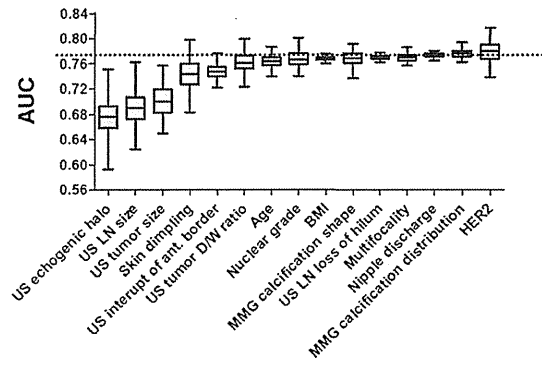


Figure 4 Sensitivity analysis using the Seoul dataset. Whisker-box plots showing 0, 25, 50, 75 and 100% (from the bottom bar to the upper bar) of the area under the curve (AUC) values when the variable was randomly replaced 200 times. The horizontal dashed line indicates the AUC value in the external validation test without any variable replacement.

0.768, respectively. However, the prediction was not perfect and there are several issues that may affect the prediction performance.

Different variations in patient variables between the training and validation datasets possibly lowered the AUC values for the external validation. There were fewer patients with AxLN metastasis in the Seoul dataset (23.6%) compared with the Tokyo (29.7%) and Kyoto (30.8%) datasets, although this was not statistically significant ($P = 0.29$) (Table 1). One reason for this difference is that patients who underwent ALND were included in the Tokyo and Kyoto datasets (14.8%) but not in the Seoul dataset. Interestingly, the number of node-positive patients in the Tokyo and Kyoto datasets was slightly higher among patients who underwent ALND compared with those who underwent SLN (39% *vs.* 29%), although this was not significant ($P = 0.15$). Despite these differences, the AUC values for the Kyoto and Seoul datasets were similar (0.770 and 0.772, respectively).

The calibration plot (Appendix D (Additional file 1)) revealed that the predictive probability for the AxLN metastasis high-risk group was overestimated in both the Kyoto and Seoul datasets. Controlled bias in the training dataset consisting of approximately 50% of AxLN-positive patients (Appendix A (Additional file 1)) likely introduced this overestimation. As demonstrated by Rouzier *et al.* [30], the calibration curves for the Seoul dataset were improved (corrected) by fitting the data to the Kyoto dataset using a polynominal function, which resulted in near-ideal lines (*i.e.*, $y = x$). Meanwhile, the calibration plots for the lower risk groups were relatively good, even without correction, for both the Kyoto and Seoul datasets.

Sensitivity analysis revealed the degree of influence of the variables in the developed model (Figure 1 and

Appendix B (Additional file 1)). In this analysis, the values of each variable were randomized (Figure 4). Of the variables causing a greater decrease in AUC values, AxLN size is directly associated with lymph node metastasis. Tumor size is used as a predictive factor in the MSKCC nomogram [6]. Echogenic halo, interruption of the anterior border the mammary gland on ultrasonography, and skin dimpling are features that reflect tumor infiltration into the surrounding tissue [31,32]. Therefore, these variables might represent tumor characteristics in the prediction models.

The mean AUC values obtained for the missing value analysis (0.884 for Kyoto and 0.688 for Seoul) were very different from those obtained for all individuals (0.770 for Kyoto and 0.772 for Seoul) because of the small number of individuals with missing values. However, the differences between the upper and lower CIs were small (0.0047 for Kyoto and 0.0081 for Seoul), which indicates that the developed model has low sensitivity to missing values. One possible reason for this feature is that ADTree can calculate a range of predictive probabilities, even for cases with missing values (see the legend of Appendix C (Additional file 1)). By contrast, standard 'if–then' decision trees and CART models cannot calculate this probability. In addition to the simple structure and high accuracy of ADTree analysis, this tolerance to the missing value is also valuable when applying machine learning to clinical data with missing values.

In the pruning analysis, the AUC values for the datasets from all three institutes generally improved according to the number of ADTrees in the prediction model (Appendix E (Additional file 1)). Although increasing the number of trees resulted in a more complex model that requires more calculation time for prediction, the model developed using the ensemble procedure showed improved accuracy and generalizability.

The AUC value of the MSKCC nomogram for the authors' own external validation sets was 0.754 [6], which is similar to our own for the Seoul dataset (0.772). Therefore, the AUC values of the developed model, the MSKCC nomogram, and the Russells Hall Hospital scoring system were compared with an external validation dataset (Seoul), which yielded values of 0.777 (95% CI: 0.689–0.864, $P < 0.001$), 0.664 (95% CI: 0.560–0.768, $P = 0.0033$) and 0.620 (95% CI: 0.509–0.731, $P = 0.0032$), respectively (Appendix F (Additional file 1)). The higher AUC value for our ADTree method might be attributed to the flexible model structure and the greater number of variables incorporated into the model. By comparison, the main advantage of both the MSKCC nomogram and the Russells Hall Hospital scoring system is that they require a small number of variables, which can facilitate data collection and interpretation of the model. Thus, these features of each modeling method represent trade-offs that should be considered when applying the models.

In addition to AUC value-based prediction performance, the false-negative rate (FNR) of the prediction model is also important when applying these models in clinical settings. For example, when a predictive value of ≤ 20% is defined as low risk for AxLN metastasis, the FNR of both the ADTree model and the MSKCC nomogram using the Seoul dataset was relatively good (5.3% and 2.6%, respectively). However, the nomogram predicted that only 6.9% of the patients were AxLN negative, compared with 23.7% using the developed model.

Unlike the MSKCC nomogram and our ADTree model, Reyal et al. developed MLR-based nomograms using the molecular subtype classification defined by a combination of ER and HER2 status with clinical parameters that included tumor size, LVI and age [33]. The decision to use ER/HER2 subtype might be attributed to the expected relationship between intrinsic breast cancer subtype and lymph node metastasis. Instead, we treated these variables as independent possible predictive factors and ADTree did not select ER status, but did select HER2 status in model development. Interestingly, HER2 status showed the lowest sensitivity in our model and the contribution of this subtype-related variable to AxLN metastasis was not significant in our study.

There are several limitations and perspectives to be discussed. First, to eliminate inter-institute or inter-interpreter variations, a standardized ultrasonography/mammography scoring system is vital because these variables are key factors for the accurate prediction of AxLN metastasis. Since a larger number of variables is required to achieve accurate prediction, unlike conventional prediction models or scoring systems, a web-based user interface, such as the one used for the MSKCC nomogram [6], will help to encourage its use and to ensure it is used correctly. In addition to calculating the probability of AxLN metastasis, a web-based platform can also assist with data collection and ensure the prediction model is kept up to date. Alternatively, machine learning-based medical classification systems have been developed following the introduction of electronic medical record systems [34-36]. Integrating prediction tools with electronic record systems will enable researchers not only to improve classification algorithms using high-dimensional datasets, but also to avoid time and effort transferring data into the classification system. Although the variables used in our developed model are frequently assessed in preoperative examinations, our proposed model is very flexible as it can incorporate new diagnostic methods or criteria. We are now developing a web-based platform to allow wider use of our model. Finally, further validation using prospective and larger datasets is indispensable before it can be used clinically.

Conclusions

In summary, we have developed a new data-mining approach based on a combination of ADTrees to predict AxLN status in patients with primary breast cancer, as a case study. The modeling method showed accurate and versatile prediction using datasets from three institutions, despite using a large number of variables. This is one of the main benefits of using data-mining methods, unlike conventional MLR methods that can only use a few independent variables to eliminate multicollinearity. The robustness of the model against missing values is also an important property of prediction models. We believe that the approach used here could replace the conventional statistical methods and provide useful information to aid decision-making before starting treatment.

Additional files

Additional file 1 Appendix A: Processes used to develop the predictive model. Additional B: ADTree-based prediction models. Additional C: Calculation of the predictive score in each ADTree model. Additional D: Calibration plots of the ADTree-based model for the Kyoto and Seoul datasets. Additional E: AUC values and the number of nodes in the pruning analysis. Additional F: ROC curves of the ADTree model, the MSKCC nomogram and the Russells Hall Hospital scoring system using the Seoul dataset ($n = 131$).

Abbreviations

AxLN: Axillary lymph node; ADTree: Alternating decision tree; ROC: Receiver operating characteristics; AUC: Area under the receiver operating characteristics curve; SLN: Sentinel lymph node; ALND: Axillary lymph node dissection; MLR: Multiple logistic regression; LVI: Lymphovascular invasion; FNR: False-negative rate; MSKCC: Memorial sloan-kettering cancer center; BMI: Body mass index; HER2: Human epidermal growth factor receptor 2; CI: Confidence interval.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We wish to thank Naoya Gomi, Kazunori Kubota, Hiroko Bando and Tomoyuki Aruga for their contributions to the grading committee. We also thank Hidetaka Furuta, Nakajima Minako and Makiko Hirose for supporting this project; Dai Kitagawa, Susumu Sekine, Tomoharu Sugie, Takayuki Ueno, Hiroyasu Yamashiro, Hiroshi Ishiguro, Wakako Tsuji, Megumi Takeuchi, Soo-Kyung Ahn and Hee-Chul Shin for help with data collection; and Shinichiro Horiguchi and Yoshiki Mikami for performing the pathological diagnoses. We thank Nicholas Smith who provided medical writing services on behalf of Edanz Group Ltd.

This study was funded by research grants from the Ministry of Health, Labour and Welfare, Japan (A study on the construction of an algorithm for multimodal therapy with biomarkers for primary breast cancer by formulation of a decision making process, led by Masakazu Toi, no. H18-3JIGAN-IPPAN-007; Reduction and lowering of recurrence risk, toxicity and pharmacoeconomic cost by prediction of efficacy for anticancer agents in breast cancer patients, led by Masakazu Toi; no. H22-GANRINSHO-IPPAN-039), research funds from Yamagata Prefectural Government and Tsuruoka City, and an International Internship Grant from the Global COE project "Center for Frontier Medicine", Kyoto University. The study was also supported by the program "Raising Proficient Oncologists" run by the Japanese Ministry of Education, Culture, Sports, Science and Technology.

Author details

[1]Department of Breast Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan. [2]Research Fellow of the Japan Society for the Promotion of Science, Tokyo, Japan. [3]Institute for Advanced Biosciences, Keio University, Yamagata, Japan. [4]Systems Biology Program, Graduate School of Media and Governance, Keio University, Kanagawa, Japan. [5]Medical Innovation Center, Kyoto University Graduate School of Medicine, Kyoto, Japan. [6]Department of Environment and Information Studies, Keio University, Kanagawa, Japan. [7]Department of Surgery, Seoul National University College of Medicine, Seoul, Republic of Korea. [8]Department of Health Care Policy and Management, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Ibaraki, Japan. [9]Department of Surgery, Tokyo Metropolitan Cancer and Infectious Diseases Center, Komagome Hospital, Tokyo, Japan. [10]Department of Pathology, Tohoku University Hospital and School of Medicine, Miyagi, Japan. [11]Department of Breast Surgery, Tenri Hospital, Nara, Japan.

Authors' contributions

MT (Takada) carried out the statistical analysis. MS performed data-mining analysis. MT, MS and YN drafted the manuscript. HM, WH and DN collected the validation data and drafted the manuscript. MK helped to design the study and helped to draft the manuscript. KK collected the training data. HS, TI and MT (Tomita) helped to design the study. MT (Toi) conceived the fundamental idea, designed the study and drafted the manuscript. All authors read and approved the final manuscript.

Received: 30 December 2011 Accepted: 13 June 2012
Published: 13 June 2012

References

1. Fisher B, Bauer M, Wickerham DL, Redmond CK, Fisher ER, Cruz AB, Foster R, Gardner B, Lerner H, Margolese R, *et al*: **Relation of number of positive axillary nodes to the prognosis of patients with primary breast cancer. An NSABP update.** *Cancer* 1983, **52**:1551–1557.
2. Fitzgibbons PL, Page DL, Weaver D, Thor AD, Allred DC, Clark GM, Ruby SG, O'Malley F, Simpson JF, Connolly JL, *et al*: **Prognostic factors in breast cancer. College of American Pathologists Consensus Statement 1999.** *Arch Pathol Lab Med* 2000, **124**:966–978.
3. Nemoto T, Natarajan N, Bedwani R, Vana J, Murphy GP: **Breast cancer in the medial half. Results of, National Survey of the American College of Surgeons.** *Cancer* 1978, **1983**(51):1333–1338.
4. Rosen PP, Groshen S, Saigo PE, Kinne DW, Hellman S: **Pathological prognostic factors in stage I (T1N0M0) and stage II (T1N1M0) breast carcinoma: a study of 644 patients with median follow-up of 18 years.** *J Clin Oncol* 1989, **7**:1239–1251.
5. Barranger E, Coutant C, Flahault A, Delpech Y, Darai E, Uzan S: **An axilla scoring system to predict non-sentinel lymph node status in breast cancer patients with sentinel lymph node involvement.** *Breast Cancer Res Treat* 2005, **91**:113–119.
6. Bevilacqua JL, Kattan MW, Fey JV, Cody HS 3rd, Borgen PI, Van Zee KJ: **Doctor, what are my chances of having a positive sentinel node? A validated nomogram for risk estimation.** *J Clin Oncol* 2007, **25**:3670–3679.
7. Chagpar AB, Scoggins CR, Martin RC 2nd, Carlson DJ, Laidley AL, El-Eid SE, McGlothin TQ, McMasters KM: **Prediction of sentinel lymph node-only disease in women with invasive breast cancer.** *Am J Surg* 2006, **192**:882–887.
8. Cho J, Han W, Lee JW, Ko E, Kang SY, Jung SY, Kim EK, Moon WK, Cho N, Park IA, *et al*: **A scoring system to predict nonsentinel lymph node status in breast cancer patients with metastatic sentinel lymph nodes: a comparison with other scoring systems.** *Ann Surg Oncol* 2008, **15**:2278–2286.
9. Degnim AC, Reynolds C, Pantvaidya G, Zakaria S, Hoskin T, Barnes S, Roberts MV, Lucas PC, Oh K, Koker M, *et al*: **Nonsentinel node metastasis in breast cancer patients: assessment of an existing and a new predictive nomogram.** *Am J Surg* 2005, **190**:543–550.
10. Hwang RF, Krishnamurthy S, Hunt KK, Mirza N, Ames FC, Feig B, Kuerer HM, Singletary SE, Babiera G, Meric F, *et al*: **Clinicopathologic factors predicting involvement of nonsentinel axillary nodes in women with breast cancer.** *Ann Surg Oncol* 2003, **10**:248–254.
11. Kohrt HE, Olshen RA, Bermas HR, Goodson WH, Wood DJ, Henry S, Rouse RV, Bailey L, Philben VJ, Dirbas FM, *et al*: **New models and online calculator for predicting non-sentinel lymph node status in sentinel lymph node positive breast cancer patients.** *BMC Cancer* 2008, **8**:66.

12. Pal A, Provenzano E, Duffy SW, Pinder SE, Purushotham AD: **A model for predicting non-sentinel lymph node metastatic disease when the sentinel lymph node is positive.** *Br J Surg* 2008, **95**:302–309.
13. Perhavec A, Perme MP, Hocevar M, Besic N, Zgajnar J: **Ljubljana nomograms for predicting the likelihood of non-sentinel lymph node metastases in breast cancer patients with a positive sentinel lymph node.** *Breast Cancer Res Treat* 2010, **119**:357–366.
14. Van Zee KJ, Manasseh DM, Bevilacqua JL, Boolbol SK, Fey JV, Tan LK, Borgen PI, Cody HS 3rd, Kattan MW: **A nomogram for predicting the likelihood of additional nodal metastases in breast cancer patients with a positive sentinel node biopsy.** *Ann Surg Oncol* 2003, **10**:1140–1151.
15. Mjolsness E, DeCoste D: **Machine learning for science: state of the art and future prospects.** *Science* 2001, **293**:2051–2055.
16. Freund Y, Mason L: **The Alternating Decision Tree Learning Algorithm.** *Proceedings of the Sixteenth International Conference on Machine Learning* 1999, :124–133.
17. Horiguchi K, Toi M, Horiguchi S, Sugimoto M, Naito Y, Hayashi Y, Ueno T, Ohno S, Funata N, Kuroi K, *et al*: **Predictive value of CD24 and CD44 for neoadjuvant chemotherapy response and prognosis in primary breast cancer patients.** *J Med Dent Sci* 2010, **57**:165–175.
18. Zhou X, Xu J, Zhao Y: **Machine learning methods for anticipating the psychological distress in patients with Alzheimer's disease.** *Australasian Physical & Engineering Sciences in Medicine* 2006, **29**:303–309.
19. Tighe P, Laduzenski S, Edwards D, Ellis N, Boezaart AP, Aygtug H: **Use of machine learning theory to predict the need for femoral nerve block following ACL repair.** *Pain Med* 2011, **12**:1566–1575.
20. Krag DN, Anderson SJ, Julian TB, Brown AM, Harlow SP, Ashikaga T, Weaver DL, Miller BJ, Jalovec LM, Frazier TG, *et al*: **Technical outcomes of sentinel-lymph-node resection and conventional axillary-lymph-node dissection in patients with clinically node-negative breast cancer: results from the NSABP B-32 randomised phase III trial.** *Lancet Oncol* 2007, **8**:881–888.
21. Cody HS 3rd, Borgen PI: **State-of-the-art approaches to sentinel node biopsy for breast cancer: study design, patient selection, technique, and quality control at Memorial Sloan-Kettering Cancer Center.** *Surg Oncol* 1999, **8**:85–91.
22. D'Orsi CJ, Mendelson EB, Ikeda DM, *et al*: *Breast Imaging Reporting and Data System: ACR BI-RADS – Breast Imaging Atlas*. VA, American College of Radiology: Reston; 2003.
23. Kim T, Giuliano AE, Lyman GH: **Lymphatic mapping and sentinel lymph node biopsy in early-stage breast carcinoma: a metaanalysis.** *Cancer* 2006, **106**:4–16.
24. Sugie T, Kassim KA, Takeuchi M, Hashimoto T, Yamagami K, Masai Y, Toi M: **A Novel Method for Sentinel Lymph Node Biopsy by Indocyanine Green Fluorescence Technique in Breast Cancer.** *Cancers* 2010, **2**:713–720.
25. Che D, Liu Q, Rasheed K, Tao X: **Decision tree and ensemble learning algorithms with their applications in bioinformatics.** *Adv Exp Med Biol* 2011, **696**:191–199.
26. Breiman L: **Bagging predictors.** *Machine learning* 1996, **24**:123–140.
27. Witten I, Frank E: *Data mining: Practical machine learning tools and techniques*. 2nd edition. San Francisco, CA: Morgan Kaufmann Publishers; 2005.
28. Carmichael AR, Aparanji K, Nightingale P, Boparai R, Stonelake PS: **A clinicopathological scoring system to select breast cancer patients for sentinel node biopsy.** *Eur J Surg Oncol* 2006, **32**:1170–1174.
29. Sharifi S, Peterson MK, Baum JK, Raza S, Schnitt SJ: **Assessment of pathologic prognostic factors in breast core needle biopsies.** *Mod Pathol* 1999, **12**:941–945.
30. Rouzier R, Pusztai L, Garbay JR, Delaloge S, Hunt KK, Hortobagyi GN, Berry D, Kuerer HM: **Development and validation of nomograms for predicting residual tumor size and the probability of successful conservative surgery with neoadjuvant chemotherapy for breast cancer.** *Cancer* 2006, **107**:1459–1466.
31. Hashimoto H, Suzuki M, Oshida M, Nagashima T, Yagata H, Shishikura T, Imanaka N, Nakajima N: **Quantitative ultrasound as a predictor of node metastases and prognosis in patients with breast cancer.** *Breast Cancer* 2000, **7**:241–246.
32. Kijima Y, Yoshinaka H, Koriyama C, Funasako Y, Natsugoe S, Aikou T: **Ultrasound examination is useful for prediction of histologic type in invasive ductal carcinoma of the breast.** *Ultrasound Med Biol* 2008, **34**:517–524.
33. Reyal F, Rouzier R, Depont-Hazelzet B, Bollet MA, Pierga JY, Alran S, Salmon RJ, Fourchotte V, Vincent-Salomon A, Sastre-Garau X, *et al*: **The molecular subtype classification is a determinant of sentinel node positivity in early breast carcinoma.** *PLoS One* 2011, **6**:e20297.
34. Meyfroidt G, Güiza F, Cottem D, Becker WD, Loon KV, Aerts JM, Berckmans D, Ramon J, Bruynooghe M, Berghe GV: **Computerized prediction of intensive care unit discharge after cardiac surgery: development and validation of a Gaussian processes model.** *BMC Med Inform Decis Mak* 2011, **11**:64.
35. Del Fiol G, Haug PJ: **Classification models for the prediction of clinicians' information needs.** *J Biomed Inform* 2009, **42**:82–89.
36. Hazlehurst B, Frost HR, Sittig DF, Stevens VJ: **MediClass: A system for detecting and classifying encounter-based clinical events in any electronic medical record.** *Journal of the American Medical Informatics Association: JAMIA* 2005, **12**:517–529.

doi:10.1186/1472-6947-12-54
Cite this article as: Takada *et al.*: Prediction of axillary lymph node metastasis in primary breast cancer patients using a decision tree-based model. *BMC Medical Informatics and Decision Making* 2012 **12**:54.

インフォームドコンセントのための図説シリーズ

乳がん薬物療法

改訂版

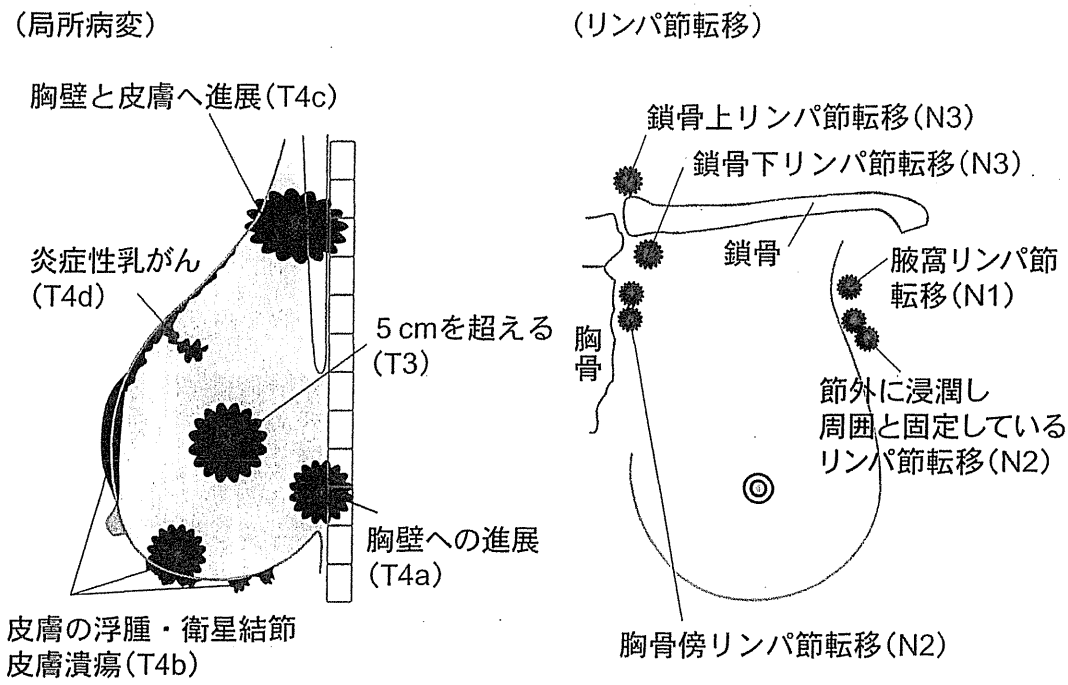
戸井 雅和 編
京都大学大学院医学研究科外科学講座乳腺外科学教授

医薬ジャーナル社

2．局所進行乳がんの治療戦略

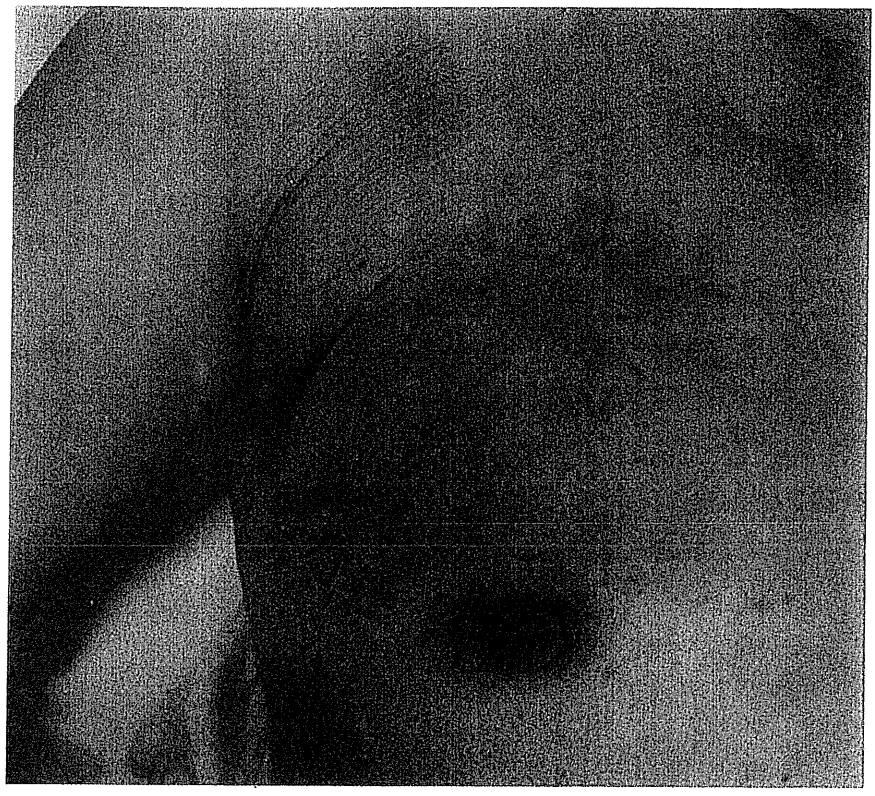
局所進行乳がんとは

『局所進行乳がん』とは乳房にできたがんが大きく，皮膚や胸壁へ浸潤している場合や，所属リンパ節への顕著な転移があるものの，画像検査で遠隔臓器への転移がみつからない状態をさします。また，乳房皮下のリンパ管にがん細胞がひろく入り込んだ結果，乳房全体が炎症を起こしたように見える炎症性乳がんも局所進行乳がんに分類されており，病期分類ではステージⅢの乳がんの総称として考えられています。ステージⅢは局所の病気のひろがりやリンパ節転移の状況などにより，ステージⅢＡ～ⅢＣに分類されています。



局所進行乳がん

参考：腫瘍径が 2 cm 以下（T1），腫瘍径が 2 cm ～ 5 cm 以下（T2）



炎症性乳がん：T4d

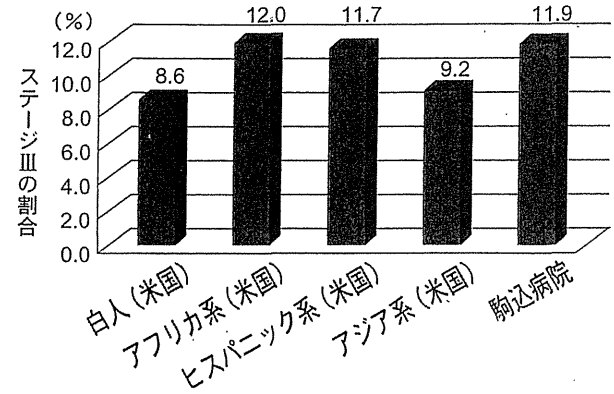
ステージⅢの乳がん

ステージ	局所病変	リンパ節転移
ⅢA	T3	N1
	Any	N2
ⅢB	T4a～d	N0，N1，N2
ⅢC	Any	N3

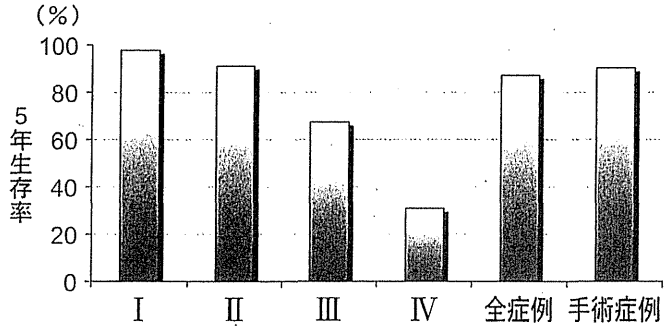
局所進行乳がんの現状

局所進行乳がんは，マンモグラフィー検診の普及などにより減少傾向にあるものの，2006年のアメリカの乳がん患者データベースにおいても9.1%を占めています。駒込病院において1998年～2010年までに治療を行った2,633人中，最初にステージⅢと診断された人は313人（11.9%）と，アメリカでの発表と大きな差はありませんでした（次ページ上図）。

また，全国がん（成人病）センター協議会のデータによれば，ステージⅢ乳がんの5年生存率は67.8%と報告されており，ステージⅠ（98.2%），ステージⅡ（91.5%）の生存率と比べて悪いことが知られています（次ページ下図）。



最初の診断でステージⅢと診断される人の割合
(Newman, et al : Seminars in Radiation Oncology 19 : 195-203, 2009 より引用改変)



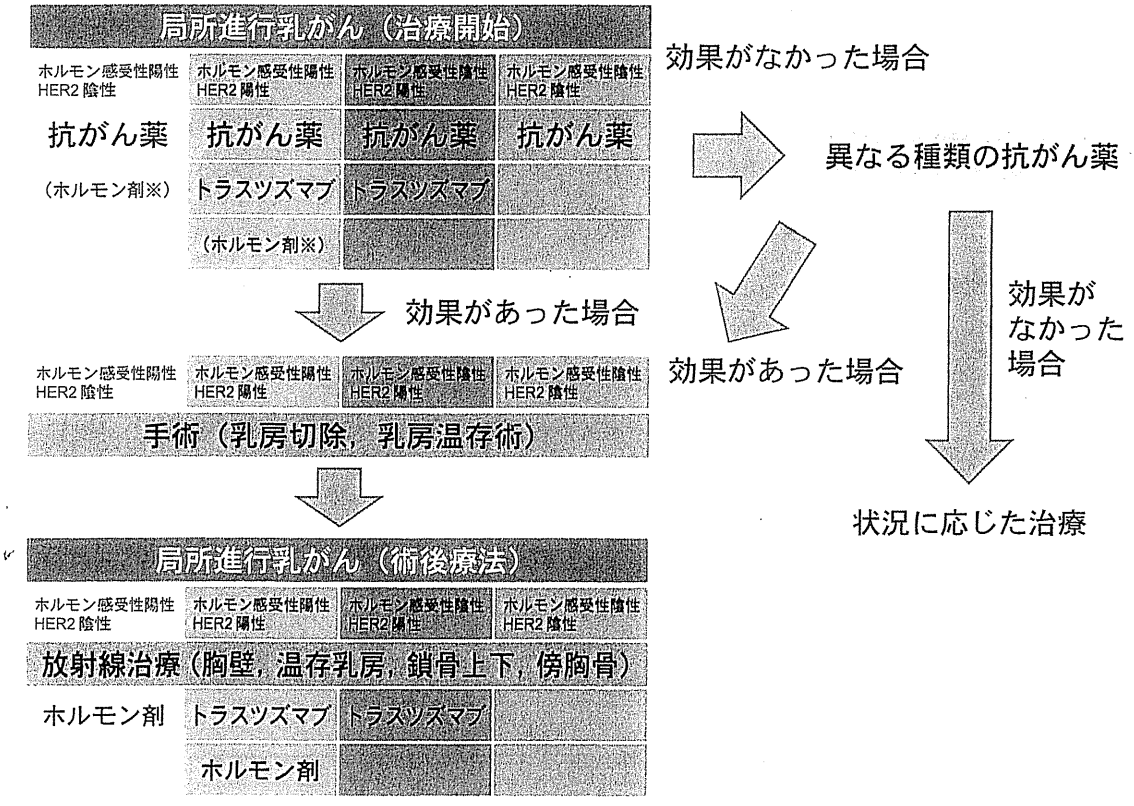
乳がんステージ別生存率
全国がん(成人病)センター協議会公表データより抜粋
(http : //www.gunma-cc.jp/sarukihan/seizonritu/zen_seizonritu/9.html)

局所進行乳がんに対する治療

局所進行乳がんにおいては治療開始前の画像検査にて，肺，肝臓，骨といった乳房から離れた重要な臓器への転移はないようにみえても，画像検査では見つけられない小さながん細胞がすでに全身にひろがってしまっていることが多いため，全身療法，局所療法を駆使して根治を目指すことが治療の基本的な考え方となります。

すなわち，乳房にできたがん（原発巣）や乳房の周りのリンパ節（所属リンパ節）へひろがったがんをコントロールする治療(局所療法：手術，放射線療法)とそこから全身に散らばってしまったかもしれないがんをコントロールする治療（全身療法：抗がん薬，ホルモン剤，抗HER2 療法など)を組み合わせて行うことが重要となります。有効な全身療法は，がんの性格（ホルモン感受性の有無，HER2 発現程度）によって選択します。

一般的には，診断された時点ですぐに手術を行うと非常に大きな手術が必要になり，取り残しなく切除することも難しいことから，先に化学療法を行い，局所病変を縮小させてから手術を行います。



治療のコンセプト

※高齢，全身状態により抗がん薬投与が安全に行いにくい場合には，選択肢としてあげられます。

局所進行乳がんにおける今後の期待

局所進行乳がんは，さまざまな治療法を組み合わせて治療を進めていく必要がありますが，それゆえに近年の治療法の進歩により，最も利益を受けている乳がんともいえます。

しかし，どのような治療をどのような順番・組み合わせで行っていくのかについては，まだまだわかっていないことも多くあるのが現状です。

このような疑問を解決すべく，局所進行乳がんに対する治験，臨床試験が数多く行われ，また現在も進行中であり，さらなる改善が期待されております。

（有賀 智之，黒井 克昌）

TOPICS2

ASCOサマリー：乳腺腫瘍分野

黒井　克昌　がん・感染症センター都立駒込病院副院長

はじめに

第48回米国臨床腫瘍学会（ASCO）年次総会（Michael P. Link 会長）が2012年６月１～５日まで米国イリノイ州シカゴ（McCormick Place）にて開催された（写真１，２）。本学会は国際学会として位置づけられており，毎年，全世界から多数の関係者が参加している。2012年は31,250人が参加し，そのうち52%が米国外からの参加であった。日本からの参加者は1,312人で，ドイツに次いで第３位であった。専門医別では腫瘍内科医が24%と最も多く（外科医は２%），関心領域では乳がんが25%とトップであった。

2012年は「がん克服のための連携」のテーマのもと，135ヵ国から5,264演題の応募があり，257演題が口演，2,487演題がポスターとして採択された（採択率52%）（表１）。一方，1,859演題が ePublication のみとなり，456演題（8.7%）がリジェクトされた。国別にみると，米国からが2,346演題（44.6%）と最も多く，日本からは第２位となる253演題（4.8%）が応募され，僅差でフランス，イタリア，ドイツが続いている。ここ３年間の推移をみると，日本からの演題数はほぼ横ばいであるが，中国からの応募がしだいに増加している。また，2012年はインドからの応募が70演題に上り，トップ12入りを果たした。領域別では乳がんが509演題と最も多く，そのうち320演題が採択され，HER2/ER と Triple negative/Cytotoxics/Local therapy に分けられ発表が行われた（採択率63%）。本稿では乳がん領域における分子標的薬の話題を中心に報告する。

1 注目された分子標的薬の第Ⅲ相臨床試験

1．EMILIA 試験

トラスツズマブ（ハーセプチン®）およびタキサン系薬剤による治療歴があり，転移に対する治療中もしくは補助療法から６ヵ月以内に増悪した HER2陽性局所進行・転移性乳がんを対象として，トラスツズマブ-emtansine（T-DM1）単独とカペシタビン＋ラパチニブ（タイケルブ®）の併用を比較した国際試験で，６月３

写真１　McCormick Place

写真２　学会会場の様子

表1　参加者，演題応募数

	2010年	2011年	2012年
総参加登録者数	32,700人	31,800人	31,250人
日本	1,279人	1,215人	1,312人
総演題応募数	5,149	4,923	5,264
米国	2,316(45.0%)	2,151(43.7%)	2,346(44.6%)
日本	246(4.8%)	227(4.6%)	253(4.8%)
韓国	121(2.3%)	74(1.5%)	105(2.0%)
中国	99(1.9%)	147(3.0%)	152(2.9%)
インド	—	—	70(1.3%)

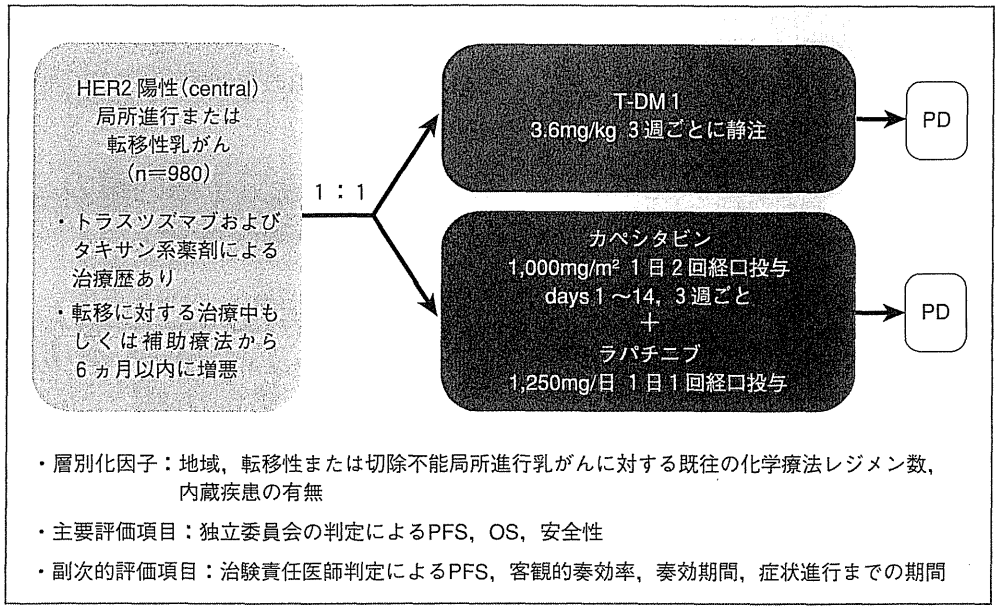


図1　EMILIA試験デザイン[1)]

日のプレナリーセッションにて結果が発表された(図1)[1)]。当日は，午後1～4時までプレナリーセッションのみが行われ，各がん種から採択された4演題のうちトップバッターとして発表された。

T-DM1はトラスツズマブに細胞障害性を有する化学療法剤DM1を結合させた抗体－薬物複合体で，HER2シグナル伝達を阻害する作用とDM1の微小管重合阻害作用により抗腫瘍効果を発揮する薬剤である。なお，DM1はHER2に結合したT-DM1が細胞内に取り込まれた後にリソソームで分解され放出される。この試験には991例が登録され，T-DM1が主要評価項目である無増悪生存期間(PFS)を有意に延長させることが示された(フォローアップ期間中央値約1年，ハザード比0.65，$p<0.0001$)(図2)[1)]。サブグループ解析の結果，65歳以上ではT-DM1群の優位性が認められなかったが，ほかのサブグループ(地域，前化学療法レジメン数，内臓転移の有無，ホルモンレセプター発現状況，治療ライン)では一貫してT-DM1群が優っていた。さらに，全生存期間(OS)も中間解析ながらT-DM1群が有意に良好であった(ハザード比0.621，$p=0.0005$)。

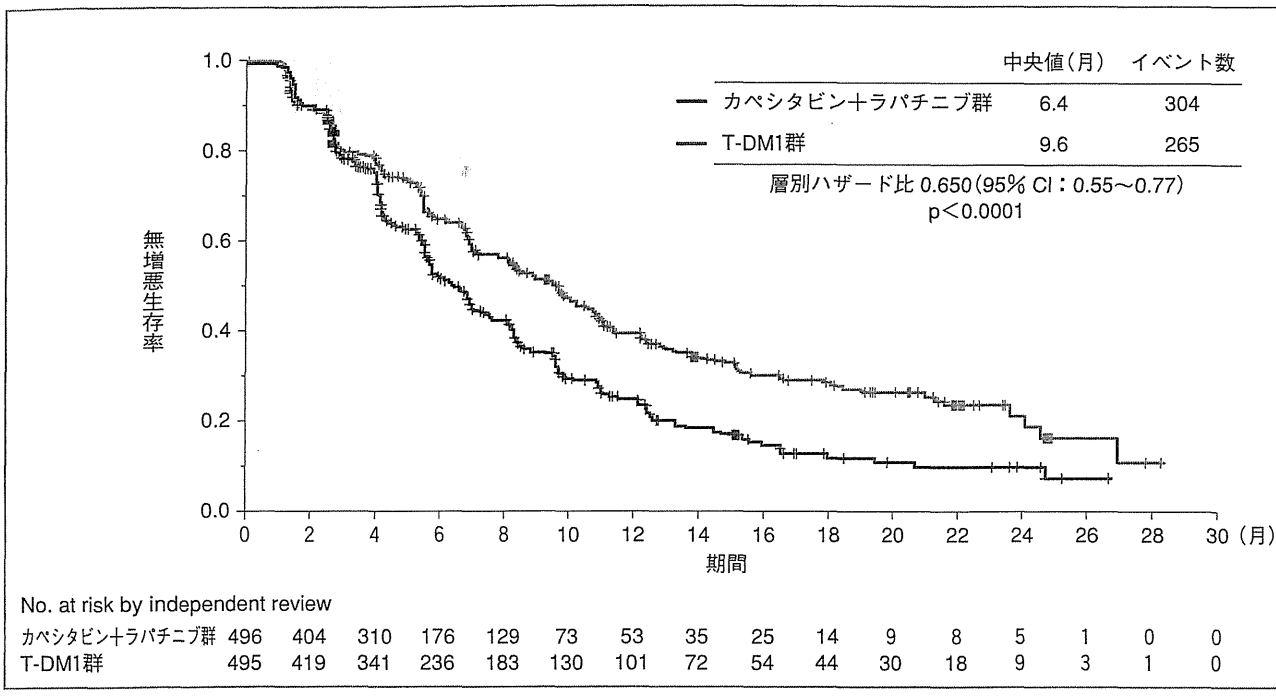


図2　独立委員会の判定による PFS(EMILIA 試験)[1]
非層別ハザード比0.66(p<0.0001)

奏効率も T-DM1群では43.6%，カペシタビン＋ラパチニブ群では30.8%で，12.7%(95%CI：6.0～19.4，p＝0.0002)の差がみられた。奏効期間の中央値は T-DM1群が12.6ヵ月(95%CI：8.4～20.8)であるのに対し，カペシタビン＋ラパチニブ群は6.5ヵ月(95%CI：5.5～7.2)であった。この試験では FACT-Breast Trial Outcome Index を用いて症状増悪(ベースラインからの5ポイント低下)までの期間が評価され，中央値は T-DM1群が7.1ヵ月，カペシタビン＋ラパチニブ群が4.6ヵ月で，T-DM1群のほうが有意に優っていた(ハザード比0.80，p＝0.0121)。

グレード3以上の有害事象は T-DM1群の40.8%，カペシタビン＋ラパチニブ群の57.0%で認められたが，これらのうち T-DM1群で多かったのは血小板数減少(グレード3　10.4%，グレード4　2.4%)，肝機能値上昇(グレード3以上；AST 4.3%，ALT 2.9%)，貧血(グレード3　2.7%，グレード4　0%)で，カペシタビン＋ラパチニブ群で多かったものは下痢(グレード3以上 20.7%)，手足症候群(グレード3以上 16.4%)，嘔吐(グレード3以上 4.5%)などであった。なお，dose intensity の中央値は T-DM1が99.9%，カペシタビンが77.2%，ラパチニブが93.4%で，それぞれ16.3%，53.4%，27.3%の症例で減量が行われていた。

2．CLEOPATRA 試験

HER2陽性転移性乳がんのファーストラインとして，ドセタキセル＋トラスツズマブ＋pertuzumab とドセタキセル＋トラスツズマブ＋プラセボを比較した国際試験で，今回は有害事象，QOL，心毒性について発表された(図3)[2]。試験治療期間は pertuzumab 併用群が18.1ヵ月，プラセボ群が11.8ヵ月であるが，両群の間でドセタキセルの投与回数(いずれも中央値8サイクル)，薬剤強度に差はなく，ほとんどの有害事象はドセタキセル終了後に減少することが示された(表2)[2]。Pertuzumab 併用群では下痢，皮疹が多く，特に，下痢はドセタキセル終了後も19%の頻度でみられたが，忍容性は両群とも良好であった。なお，発熱性好中球減少はドセタキセル併用時のみ認められたがその発現頻