Clinical Study

ONCOLOGY

Oncology 1999;57:281–286

Artificial Neural Networks Applied to Survival Prediction in Breast Cancer

M. Lundin[a] J. Lundin[a] H.B. Burke[d] S. Toikkanen[b] L. Pylkkänen[c] H. Joensuu[a]

[a]Department of Oncology, University of Helsinki, and Departments of [b]Pathology and [c]Oncology, University of Turku, Finland, and [d]Department of Medicine, New York Medical College, Valhalla, N.Y., USA

Key Words

Breast cancer · Neural networks · Survival prediction

Abstract

In this study, we evaluated the accuracy of a neural network in predicting 5-, 10- and 15-year breast-cancer-specific survival. A series of 951 breast cancer patients was divided into a training set of 651 and a validation set of 300 patients. Eight variables were entered as input to the network: tumor size, axillary nodal status, histological type, mitotic count, nuclear pleomorphism, tubule formation, tumor necrosis and age. The area under the ROC curve (AUC) was used as a measure of accuracy of the prediction models in generating survival estimates for the patients in the independent validation set. The AUC values of the neural network models for 5-, 10- and 15-year breast-cancer-specific survival were 0.909, 0.886 and 0.883, respectively.The corresponding AUC values for logistic regression were 0.897, 0.862 and 0.858. Axillary lymph node status (N0 vs. N+) predicted 5-year survival with a specificity of 71% and a sensitivity of 77%. The sensitivity of the neural network model was 91% at this specificity level. The rate of false predictions at 5 years was 82/300 for nodal status and 40/300 for the neural network. When nodal status was excluded from the neural network model, the rate of false predictions increased only to 49/300 (AUC 0.877). An artificial neural network is very accurate in the 5-, 10- and 15-year breast-cancer-specific survival prediction. The consistently high accuracy over time and the good predictive performance of a network trained without information on nodal status demonstrate that neural networks can be important tools for cancer survival prediction.

Introduction

There is a need for new prognostic systems in cancer that can integrate an expanding number of prognostic factors [1]. The possibilities to integrate new variables into existing prognostic models, such as the TNM system, are limited [2]. An optimal system would generate survival estimates for the individual patient using all the prognostic information inherent in the available patient and tumor characteristics [3].

Artificial neural networks have been successfully used for pattern recognition and survival prediction in several clinical settings [4–6]. The advantage of a neural network is the ability of the model to capture nonlinearities and complex interactions between factors [7, 8]. Trained on a number of prognostic factors, neural networks have been reported to improve the accuracy of survival prediction

KARGER

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

0030–2414/99/0574–0281$17.50/0

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

Dr. Johan Lundin
HUCH Clinical Research Institute
Haartmaninkatu 4, PO Box 105
FIN–00290 Helsinki (Finland)
Tel. +358 9 471 75395, Fax +358 9 471 75550, E-Mail johan.lundin@helsinki.fi

for patients with lung and colorectal cancer [9–11]. In patients with breast cancer, earlier studies have reported promising results for neural network models trained on the TNM variables [10].

The decreasing incidence of lymph node and distant metastases at the time of diagnosis in patients with breast cancer [12] will reduce the value of the TNM staging system in survival prediction. In this context, the primary tumor characteristics may become the major determinants in making therapy decisions and judgements regarding prognosis [12]. Attempts have been made to create models that based on the primary tumor features could predict lymph node status and identify those patients in whom an axillary nodal dissection would be indicated [13–15]. The therapeutic role of lymph node removal even in patients likely to be node positive has, however, been questioned [16]. It has been shown that accurate survival predictions can be achieved using a traditional multivariate model based on the primary tumor features only [17]. An interesting question is therefore whether a neural network could further improve the accuracy of predictions without nodal status and decrease the need for axillary dissections performed only in order to obtain prognostic information.

We evaluated the accuracy of neural network models in predicting 5-, 10- and 15-year breast-cancer-specific survival in a series of patients diagnosed with breast cancer within a defined geographic area. The variables used to construct the prognostic models in this study can be considered basic clinicopathologic characteristics and the histologic tumor features can all be assessed from routine stainings in nonspecialized institutions. The prognostic accuracy of axillary lymph node status alone and logistic regression based on the same variables as used in the neural network model were used as references. A neural network model was also constructed to assess how accurately breast-cancer-specific survival could be predicted without information on nodal status.

Patients and Methods

Patients

In order to identify all patients diagnosed with breast cancer in the city of Turku, Finland, from 1945 to 1984, we checked the files of the Finnish Cancer Registry and of the two local hospitals, the Turku University Central Hospital and the City Hospital of Turku. Histological specimens could be reviewed in 1,566 cases, and we estimated that this accounts for about 94% of all cases with female breast cancer diagnosed in the city during the time period. The cause of death was obtained from the hospital records and autopsy reports. Patients were excluded if they received adjuvant or palliative therapy, did not undergo a radical operation with axillary dissection, had intraductal or bilateral breast cancer, or distant metastases at the time of diagnosis. Of the remaining 1,050 patients, cases with missing data (n = 36) and those who died of other causes than breast cancer within 5 years were also omitted leaving 951 patients for the final analysis. These patients were randomly assigned to a train/test set of 651 and a validation set of 300 patients. For analyses at 10 and 15 years, respectively, patients who had died of intercurrent causes were excluded from the neural network and logistic regression train/test set at the corresponding time points. The percentages of patients who had died of breast cancer of those at risk in the train/test set at 5-, 10- and 15-years of follow-up were 28 (185/651), 41 (247/600) and 59% (279/474), respectively. The corresponding figures in the validation set were 29 (86/300), 43 (114/267) and 61% (130/214), respectively. The median follow-up of the whole patient series was 17 years (range 10–44 years).

Clinicopathological Variables

The distribution of the variables in the final patient series used in the analysis is shown in table 1. Histological typing and evaluation of grade components (mitotic count, nuclear pleomorphism and tubule formation) were done according to the WHO classification [18]. The tumors were classified into three histological types: ductal (not otherwise specified, includes apocrine, mixed mucinous and atypical medullary types), lobular (infiltrating lobular carcinoma with variants) and special (includes tubular, medullary, cribriform, papillary, and pure mucinous carcinomas). The number of mitoses was expressed as an average from 10 high power fields (Leitz Orthoplan, × 40 Plan objective), and nuclear pleomorphism was defined by the degree of irregularity in size, shape and staining. Tumor necrosis was graded as none, spotty, moderate or severe, but intraductal comedo necrosis was not included in its assessment. The histological parameters were determined by one pathologist (S.T.) without any knowledge of the final outcome.

Neural Network

A three-layer neural network model was constructed with a commercially available computer program using a modified cascade method together with an adaptive gradient learning rule (NeuralWorks Predict, NeuralWare, Pittsburgh, Pa., USA). The cascade mode of construction entails adding hidden nodes, one or more than one at a time, and always connecting all the previous nodes to the current node [19]. Direct connections between input and output nodes were also allowed. The adaptive gradient learning rule uses backpropagated gradient information to guide an iterative line search algorithm. Hyperbolic tangent transfer functions were used in the hidden layer and a sigmoid function in the output layer. Symmetric activation functions, such as the hyperbolic tangent, have been shown to result in faster training because the initial weights are randomized about zero [20]. Several candidate functions were tested in the output layer, and the standard sigmoid function was found to give the highest accuracy on the test set. Network output ranged from 0 to 1. From the train/test set, 30% of the patients were randomly chosen for the test set. During training the model was scored on this test set to choose between different candidate hidden nodes and to avoid overtraining. Overtraining of the network was also reduced using a weight decay method. Several candidate networks were trained and the network with the highest accuracy on the test set at 5, 10 and 15 years, respectively, was chosen for final analysis of accuracy on the independent validation set.

Statistical Analysis

The area under the receiver operating characteristic curve (AUC) was used as a measure of accuracy of the predictor models in separating survivors from nonsurvivors [21]. The statistical comparison of the areas under two ROC curves was performed according to Hanley and McNeil [22]. The difference in proportions of correctly predicted survival when comparing the neural network model to nodal status was tested for significance using the McNemar test for paired proportions. Logistic regression (SPSS for Windows, SPSS, Chicago, Ill., USA) was performed entering all variables, coded as in the neural network analysis and without interaction terms. Cox proportional hazards model was also considered as an alternative statistical model. However, we chose to present the results of the logistic regression, because the present study analyzed binary outcomes at 5, 10 and 15 years, rather than time to event data.

Results

Two final neural network models for each time point (5, 10 and 15 years) were chosen to generate survival estimates for the patients in the independent validation set, one that included axillary nodal status as an input variable and another one from which the information on nodal status had been excluded. The corresponding logistic regression survival estimates were calculated using the β-coefficients given in table 2.

The highest accuracy for predicting breast-cancer-specific survival in the validation set was achieved with the neural network models including all variables (fig. 1). When nodal status was excluded as an input variable, the accuracy of the models decreased slightly. This was also true for the logistic regression models. Both the neural network models (nodal status included and excluded, respectively) resulted in higher AUC values than the corresponding logistic regression models at all time points (fig. 1, table 3). The differences did not, however, reach statistical significance.

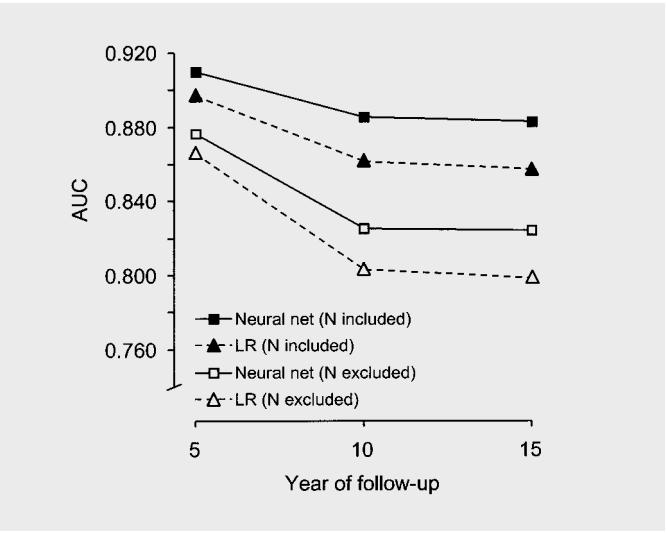
In the validation set, 66 of the 86 patients who died of breast cancer within 5 years were axillary lymph node positive, which gives a sensitivity of 77% for nodal status in predicting 5-year survival. In 152 of the 214 patients who were alive at 5 years, no positive nodes were found and thus, the specificity for nodal status was 71%. At this 71% specificity level the sensitivity of the neural network model was 90% (77/86) with all variables included and 83% (71/86) with nodal status excluded. At the same sensitivity level as that achieved by nodal status (77%), the corresponding specificities for the neural network models were 88% (189/214) and 79% (168/214), respectively.

The total rate of false survival predictions at 5 years by nodal status was 82/300. By changing the classification threshold for the survival estimates generated by the neural network model, it was possible to decrease the total number of false predictions at 5 years with nodal status included to 40/300 and with nodal status excluded to 49/300. Both rates of false predictions were significantly lower than the rate for nodal status alone (McNemar $p < 0.0001$ and $p < 0.002$, respectively). The corresponding lowest numbers of false predictions for the logistic regression models were 42/300 (nodal status included) and 55/300 (nodal status excluded).

Table 1. Distribution of clinicopathologic variables in the train/test and validation sets of patients with breast cancer

Clinicopathologic variables	Patients, n (%)		Coding
	train/test set	validation set	
Age			continuous
Median years	58	56	
Range	24–86	28–93	
Primary tumor size			
T1	210 (32)	103 (34)	1
T2	314 (48)	144 (48)	2
T3	83 (13)	27 (9)	3
T4	44 (7)	26 (9)	4
Axillary nodal status			
pN0	374 (57)	172 (57)	0
pN+	277 (43)	128 (43)	1
Histological type			categorical
Special	65 (10)	33 (11)	–1,–1
Lobular	92 (14)	40 (13)	1, 0
Ductal	494 (76)	227 (76)	0, 1
Tubule formation			
Extensive	18 (3)	14 (5)	1
Moderate	158 (24)	86 (29)	2
Slight/none	475 (73)	200 (67)	3
Nuclear pleomorphism			
Slight	103 (16)	56 (19)	1
Moderate	387 (59)	177 (59)	2
Severe	161 (25)	67 (22)	3
Mitotic count			
<2	270 (41)	133 (44)	1
2–3	236 (36)	105 (35)	2
>3/HPF	145 (22)	62 (21)	3
Tumor necrosis			
None	474 (73)	209 (70)	0
Spotty	95 (15)	50 (17)	1
Moderate	58 (9)	26 (9)	2
Severe	24 (4)	15 (5)	3

HPF = High power field.



Discussion

In this study, we have demonstrated that a neural network model trained on a number of prognostic factors can accurately predict 5-, 10- and 15-year breast-cancer-specific survival. Our results support previous reports on the prognostic performance of neural network models [6, 10, 23, 24] and confirm earlier demonstrations that have shown cancer survival predictions to be as accurate as those achieved by standard statistical models [11, 25, 26].

Fig. 1. AUCs for the neural network and logistic regression (LR) models in predicting 5-, 10- and 15-year breast-cancer-specific survival in the validation set. N = Nodal status.

Table 2. Logistic regression coefficients and significance of covariates of 5-, 10- and 15-year breast-cancer-specific outcome in the train/test set

	Nodal status included			Nodal status excluded		
	5 years	10 years	15 years	5 years	10 years	15 years
Age	−0.00	0.01	0.04*	−0.01	0.00	0.02*
Primary tumor size	0.52***	0.62***	0.47**	0.87***	0.95***	0.92***
Axillary nodal status	1.91***	1.66***	1.91***			
Histological type	*	***	***	**	***	***
Categorical coding 1	0.63	1.07*	0.58	0.71*	1.05**	0.55
Categorical coding 2	0.69*	0.91**	1.06***	0.99***	1.16***	1.24***
Tubulus formation	0.87**	0.32	0.39	0.80**	0.38	0.43
Nuclear pleomorphism	0.20	−0.17	−0.61*	0.12	−0.17	−0.52*
Mitotic count	0.59***	0.72***	0.76***	0.71***	0.76***	0.80***
Tumor necrosis	0.37*	0.29*	0.23	0.26*	0.23	0.17
Constant	−7.80***	−5.83***	−5.59***	−6.97***	−5.61***	−5.38***

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; values which are not marked by an asterisk are nonsignificant. For variable coding see table 1.

Table 3. AUC for the neural network and logistic regression (LR) models in predicting 5-, 10- and 15-year breast-cancer-specific survival in the validation set

	5 year		10 years		15 years	
	AUC	95% CI	AUC	95% CI	AUC	95% CI
Neural net (N included)	0.909	0.874–0.944	0.886	0.847–0.925	0.883	0.836–0.929
Neural net (N excluded)	0.877	0.832–0.921	0.826	0.776–0.875	0.824	0.767–0.881
LR (N included)	0.897	0.858–0.936	0.862	0.817–0.906	0.858	0.807–0.909
LR (N excluded)	0.867	0.822–0.913	0.804	0.751–0.857	0.799	0.739–0.859

N = Nodal status; CI = confidence interval.

In our study the neural networks were consistently, although not significantly, more accurate than the corresponding logistic regression models. We were also able to show that the neural network could generate accurate predictions without information on nodal status.

The high prognostic accuracy of neural network models has been ascribed to their ability to model complex relationships between prognostic variables [7, 27, 28]. As compared to traditional statistical models, neural network models can include factors that are time dependent, interact with other factors or are related to prognosis in a nonlinear fashion [7, 8]. Although statistical regression models can be used to model interactions and nonlinear relationships between variables, they require explicit search for these relationships by the model developer and may require construction of complex interaction terms [27].

A problem that remains to be resolved in survival analysis using neural networks is the handling of patients who die of an intercurrent cause or have incomplete follow-up. Several models have been suggested that could make use of the information from these patients, but none has yet been accepted as the method of choice [29–32]. The ability to include patients with incomplete follow-up or who die of an intercurrent cause would increase the number of patients available for neural network training.

A strength in the present study is that the patient series comprises nearly all breast cancer patients diagnosed within a defined geographic area during a specified time interval, and all having a minimum follow-up of 10 years. The extended time period during which data have been collected is, however, also a factor that has to be taken into account when examining the results. There have been changes in therapy regimens, in the number of screen-detected tumors and in the accuracy of TNM staging. The exclusion criteria for the final model accounted for some of these changes; patients who received adjuvant therapy were excluded, as were patients who did not undergo a radical operation with axillary dissection. A prognostic model to be used prospectively should ideally be constructed based on a patient series from a time period with as little changes in diagnostic routines, staging accuracy and therapeutic regimens as possible.

Axillary lymph node status is one of the most important prognostic factors in breast cancer. However, a much debated question is whether axillary lymph node dissection is indicated if the sole purpose is to gain prognostic information, especially in view of the decreasing incidence of lymph node metastases [12, 16]. When information on nodal status was excluded from the neural network model in the present study and the model was based solely on the histopathologic features of the primary tumor and age, it still predicted 5-year survival with a significantly higher accuracy than nodal status alone. The rate of false predictions increased only from 40 to 49 at 5 years in the validation set of 300 patients, as compared to 82 false predictions made based on axillary lymph node status (pN0 vs. pN+) as a single predictor. This shows that accurate predictions can be achieved without information on nodal status, but that the accuracy was yet slightly improved when information on nodal status was included in the model. It remains to be evaluated if the prognostic information in nodal status can be completely substituted using a combination of other variables. Alternatively, a combination of information obtained from a minimally invasive sentinel node biopsy [33] and primary tumor features could be used to achieve a high level of prognostic accuracy.

Future research should concentrate on collecting data from a more recent time period and determining new potential prognostic factors to be included in a neural network model. New prognostic factors are mainly found at the molecular-genetic level and will need prognostic systems that can capture complex interactions and nonlinear relationships to survival [8]. We are currently collecting data on a nationwide basis for patients diagnosed with breast cancer in Finland during 1991–1992 with the intent to train a neural network to be used in a prospective study. One of the goals would be to identify those patients whose prognosis after surgery is so good that no adjuvant systemic therapy will be needed.

We conclude that an artificial neural network, trained on a number of clinicopathological variables of patients with breast cancer, predicted survival with high accuracy. The consistent accuracy over time and the good predictive performance of a network trained without information on nodal status show that neural networks are valuable tools in cancer survival prediction.

Acknowledgments

This study was supported by grants from 'Finska Läkaresällskapet', the Finnish Cancer Foundation, Duodecim, the Arvid and Greta Olin Foundation and the Kurt and Doris Palander Foundation.

References

1 Clark GM, Hilsenbeck SG, Ravdin PM, De Laurentiis M, Osborne CK: Prognostic factors: Rationale and methods of analysis and integration. Breast Cancer Res Treat 1994;32:105–112.
2 Burke HB, Henson DE: The American Joint Committee on Cancer. Criteria for prognostic factors and for an enhanced prognostic system. Cancer 1993;72:3131–3135.
3 Fielding LP, Henson DE: Multiple prognostic factors and outcome analysis in patients with cancer. Communication from the American Joint Committee on Cancer. Cancer 1993;71: 2426–2429.
4 Baxt WG: Application of artificial neural networks to clinical medicine. Lancet 1995;346: 1135–1138.
5 Baxt WG, Skora J: Prospective validation of artificial neural network trained to identify acute myocardial infarction. Lancet 1996;347: 12–15.
6 Dybowski R, Weller P, Chang R, Gant V: Prediction of outcome in critically ill patients using artificial neural network synthesised by genetic algorithm. Lancet 1996;347:1146–1150.
7 Baxt WG: Complexity, chaos and human physiology: The justification for non-linear neural computational analysis. Cancer Lett 1994;77: 85–93.
8 Burke HB: Statistical analysis of complex systems in biomedicine; in Fisher D, Lenz HJ (eds): Learning from Data: AI and Statistics V. Berlin, Springer, 1996, vol 112, pp 251–258.
9 Bottaci L, Drew PJ, Hartley JE, Hadfield MB, Farouk R, Lee PW, Macintyre IM, Duthie GS, Monson JR: Artificial neural networks applied to outcome prediction for colorectal cancer patients in separate institutions. Lancet 1997; 350:469–472.
10 Burke HB, Goodman PH, Rosen DB, Henson DE, Weinstein JN, Harrell FE, Marks JR, Winchester DP, Bostwick DG: Artificial neural networks improve the accuracy of cancer survival prediction. Cancer 1997;79:857–862.
11 Jefferson MF, Pendleton N, Lucas SB, Horan MA: Comparison of a genetic algorithm neural network with logistic regression for predicting outcome after surgery for patients with non-small cell lung carcinoma. Cancer 1997;79: 1338–1342.
12 Cady B, Stone MD, Schuler JG, Thakur R, Wanner MA, Lavin PT: The new era in breast cancer. Invasion, size, and nodal involvement dramatically decreasing as a result of mammographic screening. Arch Surg 1996;131:301–308.
13 Barth A, Craig PH, Silverstein MJ: Predictors of axillary lymph node metastases in patients with T1 breast carcinoma. Cancer 1997;79: 1918–1922.
14 De Laurentiis M, Gallo C, De Placido S, Perrone F, Pettinato G, Petrella G, Carlomagno C, Panico L, Delrio P, Bianco AR: A predictive index of axillary nodal involvement in operable breast cancer. Br J Cancer 1996;73:1241–1247.
15 Ravdin PM, De Laurentiis M, Vendely T, Clark GM: Prediction of axillary lymph node status in breast cancer patients by use of prognostic indicators. J Natl Cancer Inst 1994;86: 1771–1775.
16 Fisher B, Redmond C, Fisher ER, Bauer M, Wolmark N, Wickerham DL, Deutsch M, Montague E, Margolese R, Foster R: Ten-year results of a randomized clinical trial comparing radical mastectomy and total mastectomy with or without radiation. N Engl J Med 1985;312: 674–681.
17 Menard S, Bufalino R, Rilke F, Cascinelli N, Veronesi U, Colnaghi MI: Prognosis based on primary breast carcinoma instead of pathological nodal status. Br J Cancer 1994;70:709–712.
18 World Health Organization: Histological Typing of Breast Tumours, ed 2. Geneva, World Health Organization, 1981.
19 Fahlman SE, Lebiere C: The cascade-correlation learning architecture; in Touretzky DS (ed): Advances in Neural Information Processing Systems. San Mateo, Morgan Kaufmann, 1990, vol 2, pp 524–532.
20 Brown M, An PE, Harris CJ, Wang H: How biased is your multi-layer perceptron? World Congress on Neural Networks. Portland, INNS, 1993, pp 507–511.
21 Metz CE: Basic principles of ROC analysis. Semin Nucl Med 1978;8:283–298.
22 Hanley JA, McNeil BJ: A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148:839–843.
23 Doyle HR, Dvorchik I, Mitchell S, Marino IR, Ebert FH, McMichael J, Fung JJ: Predicting outcomes after liver transplantation. A connectionist approach. Ann Surg 1994;219:408–415.
24 Ebell MH: Artificial neural networks for predicting failure to survive following in-hospital cardiopulmonary resuscitation. J Fam Pract 1993;36:297–303.
25 Bryce TJ, Dewhirst MW, Floyd CE Jr, Hars V, Brizel DM: Artificial neural network model of survival in patients treated with irradiation with and without concurrent chemotherapy for advanced carcinoma of the head and neck. Int J Radiat Oncol Biol Phys 1998;41:339–345.
26 Ravdin PM, Clark GM, Hilsenbeck SG, Owens MA, Vendely P, Pandian MR, McGuire WL: A demonstration that breast cancer recurrence can be predicted by neural network analysis. Breast Cancer Res Treat 1992;21:47–53.
27 Tu JV: Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol 1996;49:1225–1231.
28 White H: Learning in artificial neural networks. A statistical perspective. Neural Comput 1989;1:425–464.
29 De Laurentiis M, Ravdin PM: A technique for using neural network analysis to perform survival analysis of censored data. Cancer Lett 1994;77:127–138.
30 Faraggi D, Simon R: A neural network model for survival data. Stat Med 1995;14:73–82.
31 Liestol K, Andersen PK, Andersen U: Survival analysis and neural nets. Stat Med 1994;13: 1189–1200.
32 Ohno-Machado L, Walker MG, Musen MA: Hierarchical neural networks for survival analysis. Medinfo 1995;8:828–832.
33 Veronesi U, Paganelli G, Galimberti V, Viale G, Zurrida S, Bedoni M, Costa A, de Cicco C, Geraghty JG, Luini A, Sacchini V, Veronesi P: Sentinel-node biopsy to avoid axillary dissection in breast cancer with clinically negative lymph-nodes. Lancet 1997;349:1864–1867.