Original Article

Can Biomarkers Improve Ability of NPI in Risk Prediction? A Decision Tree Model Analysis

Baneshi MR[1, 2], Warner P[2], Anderson N[2], Tovey S[3], Edwards J[3], Bartlett JMS[4]

Abstract

Background: The Nottingham Prognostic Index (NPI) is widely-used in the UK for risk stratification of breast cancer patients. This paper aims to evaluate the ability of this index to detect patients with sufficiently low risk of recurrence that they could be spared harsh treatments, and to construct an enhanced prognostic rule that integrates biomarkers with clinical variables to achieve better risk stratification.

Methods: We undertook review of published studies of outcomes in risk groups derived by applying NPI, and report estimated event-free rates extracted from papers found. Then we analysed biological and clinical variables for 401 ER+ patients, to develop a Tree-based Survival Model (TSM), for risk prediction, and estimated event-free rates by resulting risk-groups, Kaplan-Meier (K-M) curves corresponding to TSM and NPI were plotted.

Results: We concluded that NPI does not distinguish low risk patients with a sufficiently high event-free rate to make it likely clinicians would decide treatments with potential harmful side effects can be avoided in that group. On the other hand, in the decision tree constructed, utilising 3 biomarkers, nodal status and tumour size, the 4 risk groups were clearly diverged in terms of event-free rates.

Conclusion: There is considerable potential for improved prognostic modelling by incorporation of biological variables into risk prediction. Whilst low risk patients identified by our TSM model could potentially avoid systemic treatment, higher risk patients might require additional treatment, including chemotherapy or other adjuvant treatment options. However, the decision tree model needs to be validated in a larger clinical trial cohort.

Please cite this article as: Baneshi MR, Warner P, Anderson N, Tovey S, Edwards J, Bartlett JMS. Can Biomarkers Improve Ability of NPI in Risk Prediction? A Decision Tree Model Analysis.Iran J Cancer Prev ,2010 ; Vol3, No2, p.62-74.

Keywords: Breast neoplasm, Tissue microarray data, NPI, Tree-based survival methods, Missing data

1. Health School, Kerman Medical University, Department of Biostatistics and Epidemiology, Kerman, Iran
2. Centre for Population Health Sciences, University of Edinburgh, Teviot Place, Edinburgh, UK
3. Section of Surgical and Translational Sciences, Division of Cancer Sciences and Molecular Pathology, Glasgow Royal Infirmary, Glasgow, UK
4. Endocrine Cancer Group, University of Edinburgh, Edinburgh Cancer Research Centre, Western General Hospital, Crewe Road South, Edinburgh, UK

Corresponding Author:
Mohammad Reza Baneshi PhD
Tel: (+98) 341-32 05 08 7
Email: rbaneshi@yahoo.com

Received: 23 Oct.2009
Accepted: 19 Nov.2009
IJCP 2010; 2: 62-74

Introduction

Cancer is one of the most major health problems worldwide. In UK, more than a quarter of a million new cancer cases are diagnosed per year. Breast carcinoma is the most prevalent malignancy, with one million newly diagnosed cases annually, comprising 18% of all female cancers [1]. Each year in UK, more than 44,000 women are diagnosed with breast cancer and 2400 die from it [2].

Management of patients suffering from cancer is guided using prognostic models. Prognostic models combine key patient characteristics (i.e. risk factors) to predict clinical outcomes such as recurrence of cancer. Such models are valuable tools for selection of treatment strategy and for providing a patient with information about her likely outcome[3].They are also useful for investigating the contribution of variables to the disease course, and to inform design of future studies [3].

In the case of breast cancer, Nottingham Prognostic Index (NPI) was devised to estimate the risk of recurrence and to classify patients into risk groups [4]. To do so, multiple conventional prognostic candidate variables (9 variables in total) have been investigated by applying the Cox regression model to derive an index of risk of disease recurrence [4]. This model (NPI) uses information on 3 clinical variables (number of positive nodes, tumour grade, and tumour size) to stratify into 3 risk groups, and has been widely validated [5-7]. However, it has been commented that, for a prognostic model such as NPI to be clinically useful in practice, it should be

able to identify a subset of patients with a prognosis so good that it would be safe to forego the risks of adjuvant therapy with potential harmful side effects, such as chemotherapy [8]. Furthermore, it has been argued that NPI is not capable of identifying such a subgroup [7].

NPI calculates risk based on only the clinical presentation of the tumour, at biopsy, yet there is clear evidence that breast cancer is a heterogeneous disease which includes different subtypes [9]. As an example, Estrogen and Progesterone hormonal Receptors (ER and PR), present in nearly two thirds of breast cancer specimens, are associated with growth of cancer cells, and the HER2 gene controls growth, division, and the repair of cells (10). It has been found that estimated 6-year Disease Free Survival (DFS) differed markedly between (ER+, PR+) and (ER-, PR-) cases (85% versus only 72%) [10]. However, recent developments in the field of cancer biology, have led to the development of Tissue Microarray Analysis (TMA), which enables measurement of many cancer tissue characteristics describing the underlying biology of breast cancers, and offering potential insights to likely disease course, and hence management of breast cancer patients.

TMA has been taken up by cancer researchers world-wide [11]. The tumour tissue is embedded in paraffin, and then using a hollow needle 6 millimetre diameter, tissue cores are removed from the region of interest, a process similar to clinical biopsies. Tissue cores are inserted in a new paraffin block and microtome used to cut sections for staining and/or microscopic examination. While these TMA variables could have the potential to enhance risk prediction, a challenge arises for prognostic modelling methods since TMA analysis results in many additional potential prognostic markers (i.e. biomarkers). Given the generally accepted requirement for at least 10 events per independent variable being tested, in order to develop a reliable Cox regression model [12], inclusion of TMA variables can substantially increase the size of cohort needed for regression modelling. However, as an alternative to traditional Cox regression model, Tree-based Survival Model (TSM) can be applied, since this place no limits on the number of variables included [13].

The aim of this paper is to review the ability of the NPI index to detect low risk patients and, using cohort data available for analysis. We also will integrate both clinical and biomarker (TMA) variables, through TSM analysis, so as to develop an enhanced prognostic model for breast cancer recurrence.

Material and Methods

Literature review of ability of NPI to identify 'low risk' patients

To evaluate the ability of NPI in risk stratification, in Oct 2008 using Pubmed data base, the word 'Nottingham' in the title of the paper was searched. References cited in papers which reported application of NPI, were also checked. Only papers which reported the application of the NPI to stratify patients into risk groups were considered. Papers cited by those papers selected were also checked. In the case of no report of event-free rates, if survival curves were presented, then figures were read off as accurately as possible from the plots.

Sample

A total of 401 Estrogen Receptor positive (ER+) women diagnosed during 1983 and 1999 at Glasgow Royal Infirmary formed our sample. Median follow-up time was 6.16 years and all patients received Tamoxifen for some of the follow-up time (median of 5 years). At the end of the follow-up there had been 112 recurrences. All patients were treated by surgery with curative intent and received Tamoxifen after surgery; 73 (18%) were aged under 50 years of age at diagnosis [14-16].

Variables and outcome

The main outcome of the study was Recurrence Free Survival (RFS). Data for a large number of tissue microarray variables (72 variables describing 41 protein biomarkers) and 3 clinical variables (nodal status, grade (Bloom and Richardson), and tumour size) were available. These 72 biomarkers belonged to RAS, AKT, PgR, MAPK, MTOR, BAD, and HER families [17]. Staining was analysed separately for membrane, cytoplasmic and nuclear localisation of biomarkers:

1. Membrane expression was analysed for: p118ERα, p167ERα, EGFr, HER2, phosphoHER2, HER3 (m), HER4-ICD (intracellular domain), HER4 ECD (extracellular domain)
2. Cytoplasmic expression was analysed for: ER α, ER β, p118ER α, p167ER α, phosphoHER2, HER 3, HER4-ICD, HER 4 ECD, h RAS, n RAS, k RAS, RAF 1, p259-RAF 1, p338-RAF 1, r Kip, TES, AKT 1, AKT 2, AKT 3, panAKT, p473AKT, p308AKT, mTOR, phospho-m TOR, p389-p70S6k, Tace, Tacep, MAPK, phosphor MAPK, PTEN, Bcl2, Bax, Bad, p112-Bad, Bcl-xl.
3. Nuclear expression was analysed for: ER α, ER β, Pg R, p118ER α, p167ER α, phosphoHER 2, HER 3, HER4-ICD, HER 4 ECD, hRAS, n RAS,

k RAS, RAF1, p259-RAF 1, p338-RAF1, r Kip, TES, AKT1, pan AKT, p473 AKT, p308 AKT, MAP K, phosphor MAP K, PTEN, AIB1.

In addition the gene amplification and the copy number for HER2 and AIB1 and TUNEL analysis of apoptosis were analysed.

TSM analysis

Construction of TSM

TSM has no upper limit for number of variables included. TSM involves successive binary partitioning, to sub-classify subjects into a number of smaller groups (known as nodes), that are homogeneous with respect to recurrence rates. To start, the Log-Rank test is applied to every possible cut point for each prognostic variable, so as to select the split that ensures the greatest difference in recurrence rates between the two resulting subgroups (as judged by the lowest P-value, and highest Hazard Ratio (HR)) [18]. The process then proceeds in the same way on each of the two subgroups (each parent node is split into two subgroups known as child nodes), on these child nodes, and so on. Once a sub-group cannot be sub-divided further (that is, when all splits have Log-Rank p-value > 0.05 , or the only split with P value ≤ 0.05 has its cut-point in the outer 20% of the distribution for that biomarker in the entire original sample [18], or would yield a sub-subgroup with n <30), then partitioning ceases and the un-split subgroup is declared a terminal node. The whole iterative process of sub-division creates a tree structure.

Patients with missing data are not used in the initial TSM analysis, but once each node has been selected, patients with missing data on the variable used, are assigned to an appropriate node by means of the 'surrogate variable' approach. This involves re-applying the partitioning algorithm explained above to all values for all other variables to select the second best variable cut off to achieve that split. If the surrogate variable has a missing value for that patient, then the next potential surrogate variable is tried, and so on [13].

Refinement of tree

A systematic search across all possible values, to detect the optimal split, can lead to the selection of over-optimistic and hence unstable cut points, due to the multiple testing undertaken [19]. To tackle potential over fitting, branches with P-value exceeding 0.002 (corresponding to 0.05 if a single test had been applied [20]) were revoked (including all sub-branches) [21,22]. Furthermore, to avoid groups with small numbers of patients, no split at outer 20% of distribution of biomarkers was applied [18].

Amalgamation of groups with similar survival curve

Although TSM ensures that the two terminal nodes within a branch are significantly different, it remains possible that terminal subgroups from distinct branches might have very similar survival curves. In accordance with usual practice in TSM modelling, terminal nodes will comprise a minimum of 30 patients. However, the higher the number of patients in node, the more robust the estimated event-free rate is. Although it would be possible, to decide in advance to require a higher minimum number of patients per terminal node, this leads to risk of a tree with very few 'branches', and hence of missing of high order interactions, in particular when the sample size is low. Therefore, further examination is required to rationalise the number of terminal nodes (risk groups), by examining the survival curves and event-free rates in the terminal nodes, and also considering number of patients in each node. We plotted the Kaplan-Meier survival curves and estimated the actuarial 7-year RFS rates (see below) [23,24].

Assignment to NPI risk groups

The NPI risk scores are calculated as: NPI = 0.2 x Size (cm) + Nodal status + Grade where both nodal status and grade are scored as 1, 2 or 3. Tumour size was based on measurement of the mastectomy specimen. Histological grade (1 to 3) was determined based on criteria of Bloom and Richardson [25]. The Bloom-Richardson grading method is based on three features of invasive breast cancers: the percentage cancer composed of tubular structures, the rate of cell division, and the nuclear pleomorphism of tumor cells (nuclear grade, change in cell size and uniformity). Each of these 3 features is rated from 1 to 3. Summation of these scores, which give a total score that ranges from 3 to 9, is used to grade the tumours as follows:

- Grade 1 tumor (well-differentiated): scores 3 to 5
- Grade 2 tumor (moderately-differentiated): scores 6 to 7
- Grade 3 tumor (poorly-differentiated): scores 8 to 9

Lymph node involvement was determined based on biopsy of a lower axillary node, an apical axillary node, and a node from the internal mammary chain. Patients were staged into 3 groups in terms of lymph node findings:

- Stage 1: Tumour absent from all 3 nodes sampled

- Stage 2: Tumour in low axillary node only.
- Stage 3: Tumour in either of apical or internal mammary nodes

In the data set used there was some missing data for clinical variables;altogether 58 patients had one or more missing values for clinical variables - 33 for nodal status, 11 for grade and 22 for tumour size.

In calculation of the NPI, Multivariable Imputations by Chained Equations (MICE) method [26,27] was applied to deal with missing data. The MICE method is a probability-based simulation technique which takes into account imputation uncertainty pertaining. This is an iterative process where missing data for a variable is estimated using its imputation model and then in turn these data are used in estimation of missing data for another variable. In accordance with the usual practice, we imputed 10 values for each missing value, and thus created 10 imputed data sets.

For each woman, an NPI risk score was calculated as above for each of the 10 imputed data sets, and her final NPI score was the average of these 10 risk scores. To create the risk groups, cut offs were applied to final (average) NPI risk scores, so that patients with average score ≤ 3.4 and > 5.4 formed the lowest and highest risk groups respectively, and the remainder the intermediate risk group. In this way, every patient was categorised into one of the 3 risk groups.

Kaplan-Meier (K-M) curves for TSM and NPI risk groups

K-M survival curves were plotted and actuarial 5 and 7 year event-free rates years (with 67% and 40% follow-up data respectively) were calculated corresponding to TSM and NPI risk groupings. Being event free all the way to the end of 7th year depends on no event in any of the preceding years, and also none in the 7th year. In actuarial life-table procedure, the whole follow-up duration is split into one year intervals. If indicates the number of patients at risk just before the i-th year starts, and the number of events during the i-th year, then the probability of being event free up to and including 7th year is given by

$$S(7) = \prod_{i=1}^{7} (1 - \frac{d_i}{n_i})$$

Results

Review of event free survival by NPI risk group

Our literature review resulted in 470 papers; the majority of them were not relevant to breast cancer. Only 17 papers were relevant, but two studies split each of the three risk groups into two, thus dividing the patients into 6 risk groups [28, 29]. Results of these 2 studies could not be compared with other studies since different cut offs were applied. Of the 15 papers included in the review, not all provided information on confidence interval of reported survival rates, detailed information about the number of patients and recurrences in risk groups, follow-up time, and estimated event-free rates.

A total of 15 papers reported short-term rates at 5 years (Table 1). Actuarial 5-year survival rate derived from original NPI was 88% [4]. In the literature, the estimated/ reported rate varied from 82% [30, 31] to 96% [32, 33]. In the case of 3 studies, breast cancer patients were all node negative, so none was assigned into the high risk group [30, 31, 34].

Focusing on longer-term survival rates (Table 2), in the largest studies, nearly 25000 and 10000 patients were recruited [7,35]. In both studies, estimated 10-year survival rate was about 80%. Some other studies reported a similar rate [6, 35-37].

The highest 10-year event-free rate in the low risk group was 88%, reported for a study recruiting only patients with small primary breast cancer [33]. However the same 10-year event-free rate (88 %,) was reported in the longest follow-up study [38], and the event-free rate for the high-risk group in the latter study was also higher than most of the other studies.

The poorest 10-year survival rate in a low risk group was only 66% [39]. Sample size and duration of follow-up was not reported. Callagy et al. (40) reported an estimate only slightly better (73% versus 66%).

Two cohorts were analysed in which lowest-risk patients were defined as those with NPI≤ 2.4 (not shown in the tables) [28,29]. The cohort with longer follow-up data gave 10-year survival rate of 88% [29]. The corresponding rate for the other cohort was 96% [28].

TSM and NPI Model

The tree constructed is given in Figure 1. Branching points are shown as rectangles, and numbers in square brackets are, respectively, number of patients at this branching point, and number who had recurrences. Terminal nodes are shown as ovals, with numbers as for branching points. For each split, the P-value reported corresponds to the Log-Rank P-value test. A total of 5 variables were used to construct the tree. The first two variables which best separated the patients were nodal status and tumour size.

Biological variables, Tunel, Prhisto, and cytoplasmic KRAS were also required.

The absolute difference between estimated 7-year RFS rates of nodes 3 and 4 was 20% (Table 3). Comparing nodes 1 with 2, and 5 with 6 (Table 3), rates within each pair were fairly similar (difference 6% in each comparison). Furthermore, the number of patients comprising nodes 1, 2, 5 and 6 were smaller than for nodes 3 and 4. Therefore to ensure more robust estimates nodes 1 and 2 were combined to create a single lowest risk group. In addition, nodes 5 and 6 were grouped to create a highest risk groups.

In the lowest risk group identified by TSM (nodes 1 and 2 combined), there were only 3 recurrences out of 84 patients. This gave a 7-year RFS of 96% (95% C.I. 92% to 100%). Figure 2 presents K-M curves showing survival experience for NPI risk groups and for risk groups derived by applying TSM model. While in NPI the three groups are fairly well diverged, the survival curve for the lowest risk group (top curve) continues to decline across the entire follow up period. In contrast, for the four TSM risk groups, the lowest risk group (top curve) shows very few events and no event at all after about 3 years follow-up.

Discussion

In our sample, actuarial 5 and 10-year RFS rates in the lowest NPI risk group were 94% and 79% respectively, although it should be noted that sparse follow up to the 10th year (about 11%), means that our estimated 10-year rate might not be robust. The literature review of published studies (Tables 1 and 2) showed that while the median of the short-term (5 year) recurrence free rate of the lowest-risk group defined by NPI was 90%, a gentle decreasing trend was seen in K-M curves after fifth year of follow up to a median 10-year event-free rate of 80% (ranging from 66% (39) to 88%) [33,38] Although patients were stratified by the same prognostic index (i.e. NPI), marked differences were seen between estimations. However, comparison of results is not straightforward since different patient subgroups received a variety of treatment regimes. In addition, the sizes of studies varied hugely (from 82 to more than 25000). However, our aim was neither to compare the results nor to do a quality assessment of estimations. We simply aimed to perform a narrative review, across the spectrum of study types in which NPI has been utilised, to investigate Balslev and co-authors' statement that NPI is not able to identify patients with very low risk [7].

The second aim of our research was to develop a TSM prognostic model for RFS combining clinical and biomarker variables, that could identify a clearly low risk group, that might be spared adjuvant treatments. Prior to the start of this research, this issue was discussed with the clinical collaborators of the study and it was proposed that a minimum 10-year RFS of 95% would define such a group. Elsewhere it has been commented that a prognostic group with a 15-year survival of 94% could be considered 'a group of patients potentially cured by locoregional treatment alone' [33]. Our literature review showed that the best published estimate using standard NPI is 88% 10-year RFS in the lowest risk group. Our review therefore showed that standard NPI is not capable of identifying a subgroup of patients with sufficiently low risk of recurrence who do not require harsh treatments [7].

With regard to the TSM model developed, the lowest risk group of the Glasgow data set had both 5 and 7-year RFS of 96% (95% CI: 92%, 100%). In contrast, 7-year rates for NPI were 89% (95% CI: 83%, 95%). However, an adaptation of NPI, applying an additional cut-off at ≤2.4, has identified a 'lowest risk' group with a 15-year survival of 94% [33]. However, this is essentially all those with nodal status, grade score of 1 and tumour size of ≤ 2 cm, and only a small proportion of patients are likely to meet these criteria. In the Glasgow data set this subgroup comprises only 43 patients (11%), who had an estimated 7-year, RFS of 95% (95 % CI: 89%, 100%).

Although the lowest risk group detected by TSM (n=84) had higher (better) 7-year RFS rate than the lowest risk group by NPI (n=133), it should be noted that for the same data set, event-free rates estimated for patients in the "lowest risk group", to a great extent depend on decisions as to the size of that risk group. The 84 patients with lowest NPI scores would have better RFS than the larger subgroup of 133 classified to as its 'low risk' group (by its pre-defined classification rule). So we are not comparing like with like. An alternative strategy for improving low risk group stratification in NPI would be to add a further cut off to is classification, and thus create an even lower risk NPI group, as has been done [33], or to use NPI scores to create a new grouping with more than 3 groups [41].

The main weaknesses of the TSM approach are multiple testing and potential over-fitting [19]. It has been discussed that decision trees are sensitive to small changes in the sample [42]. Sensitivity (of results) to sample is a problem in all modelling approaches (including regression methods).

Table 1. Comparison of 5-year event frees rate across studies in the subset of patients identifies as being low risk by NPI

Study	Ref. number	Cohort size	Treatment Given	Follow-up (years)	Lowest-risk group (L)		Intermediate-risk group (I)		Highest-risk group (H)	
					Number (%) of cases	**Event-free rate (%)**	Number (%) of cases	**Event-free rate (%)**	Number (%) of cases	**Event-free rate(%)**
Haybittle (1982)	(4)	387	Simple mastectomy + triple node biopsy	1- 6	64 (21)	**88**	169 (57)	**69**	65 (22)	**21**
Todd (1987)	(5)	387[1]+ 320	Simple mastectomy + triple node biopsy Simple Mastectomy	6-11.5 1.7- 6.5	192 (27)	**88**	381 (54)	**69**	134 (19)	**22**
Okugawa (2005)	(32)	311	ER- patients treated with adjuvant systematic chemotherapy consisting of cyclophosphamide, methotrexate, and 5-fluorouracil (CMF) for 5 months ER+ treated with tamoxifen for 2 years.		97 (31)	**96**	142 (46)	**85**	72 (23)	**45**
Sauerbrei (1997)	(30)	603	Modified radical mastectomy + en bloc axillary dissection with ≥ 6 identifiable lymph nodes in the specimen	5	163 (27)	**82**	440 (73)	**70**	No case	
Coradini (2001)	(31)	226	Radical or conservative surgery + radiotherapy, and complete axillary dissection until relapse	0.3- 8.17 Median 6.25		**82***		**72***	No case	
Ring (2006)	(34)	195	Only ER+ lymph node – patients were analysed. Treatments applied not given			**90**		**90**	No case	

1. These 387 patients were those used to devise the NPI index
* inexact read off from graph

Table 1 (continued). Comparison of 5-year event frees rate across studies in the subset of patients identifies as being low risk by NPI

Study	Ref. number	Cohort size	Treatment given	Follow-up (years)	Lowest-risk group (L)		Intermediate-risk group (I)		Highest-risk group (H)	
					Number (%) of cases	**Event-free rate (%)**	Number (%) of cases	Event-free rate (%)	Number (%) of cases	Event-free rate (%)
Balslev (1994)	(7)	9149	94.8% simple mastectomy 5.2% tumorectomy + radiotherapy	2.3- 13.9 median 7.1	2494 (27)	**92**	5245 (57)	75	1410 (16)	40
Kollias (1999)	(33)	2684	69.1% mastectomy or subcutaneous mastectomy 30.9% lumpectomy (more details in the paper)		894 (33)	**96***	1374 (52)	82*	416 (15)	35*
Sidoni (2004)	(36)	82	Not reported	Min 5	27 (33)	**92***	39 (48)	63*	16 (19)	55*
Eden (2004)	(37)	97[1]	5 patients adjuvant systemic therapy (no more information was provided)			**83***				43*
Callagy (2006)	(40)	557	Chemotherapy	0.4- 39.4 Median 8.7	34 (6)	**83***	236 (42)	75*	287 (52)	55*
Lundin (2006)	(35)	2036 (FinProg series)	Node negative / Node positive	Median 9.5		**90***		80*		50*
			Adjuvant therapy 8.8% 92.3% Chemotherapy 6.2% 52.0% Hormone therapy 2.4% 36.1% Unknown 0.2% 3.1%							
		25752 (SEER series)	Approximately 66% adjuvant therapy, 18% chemotherapy, 35% hormone therapy, and 9% both	Median 9.7		**91***		83*		60*
D'Eredita (2001)	(38)	402	Surgery + axillary clearance (details provided in the paper)	11-19 median 15	110 (27)	**93***	198 (49)	75*	94 (23)	50*
Galea [1] (1992)	(6)	1629	Not reported		470 (29)	**92***	879 (54)	72*	280 (17)	25*

2. Eden et al. studied 46 patients who developed distant metastasis within 5 years and 51 patients being distant metastasis-free for ≥ 5 years
3. Galea et al. reported 15-year event-free rates
* inexact read off from graph

Table 2. Comparison of long-term event free rate (10 years or more) across studies in the subset of low risk patients identifies as being low risk by NPI

Study	Ref. number	Cohort size	Treatment given	Follow-up (years)	Lowest-risk group (L)		Intermediate-risk group (I)		Highest-risk group (H)	
					Number (%) of cases	**Event-free rate (%)**	Number (%) of cases	Event-free rate(%)	Number (%) of cases	Event-free rate(%)
Brown (1993)	(39)		Not reported			**66**		50		34
Balslev (1994)	(7)	9149	94.8% simple mastectomy 5.2% tumorectomy + radiotherapy	2.3- 13.9 median 7.1	2494 (27)	**79**	5245 (57)	56	1410 (16)	25
Kollias (1999)	(33)	2684	69.1% mastectomy or subcutaneous mastectomy 30.9% lumpectomy (more details in the paper)		894 (33)	**88***	1374 (52)	58*	416 (15)	17*
Sidoni (2004)	(36)	82	Not reported	Min 5	27 (33)	**83***	39 (48)	60*	16 (19)	42*
Eden (2004)	(37)	97[1]	5 patients adjuvant systemic therapy (no more information was provided)			**83***				43*
Callagy (2006)	(40)	557	Chemotherapy	0.4- 39.4 Median 8.7	34 (6)	**73***	236 (42)	60*	287 (52)	38*
Lundin (2006)	(35)	2923 (FinProg series)	Node negative / Node positive	Median 9.5		**79***		70*		29*
			Adjuvant therapy 8.8% 92.3%							
			Chemotherapy 6.2% 52.0%							
			Hormone therapy 2.4% 36.1%							
			Unknown 0.2% 3.1%							
		25752 (SEER series)	Approximately 66% adjuvant therapy, 18% chemotherapy, 35% hormone therapy, and 9% both	Median 9.7		**80***		70%*		29*
D'Eredita (2001)	(38)	402	Surgery + axillary clearance (details provided in the paper)	11-19 median 15	110 (27)	**88***	198 (49)	70%*	94 (23)	40*
Galea [1] (1992)	(6)	1629	Not reported		470 (29)	**80***	879 (54)	42*	280 (17)	13*

4. Eden *et al.* studied 46 patients who developed distant metastasis within 5 years and 51 patients being distant metastasis-free for ≥ 5 years
5. Galea *et al.* reported 15-year event-free rates
* Inexact read off from graph

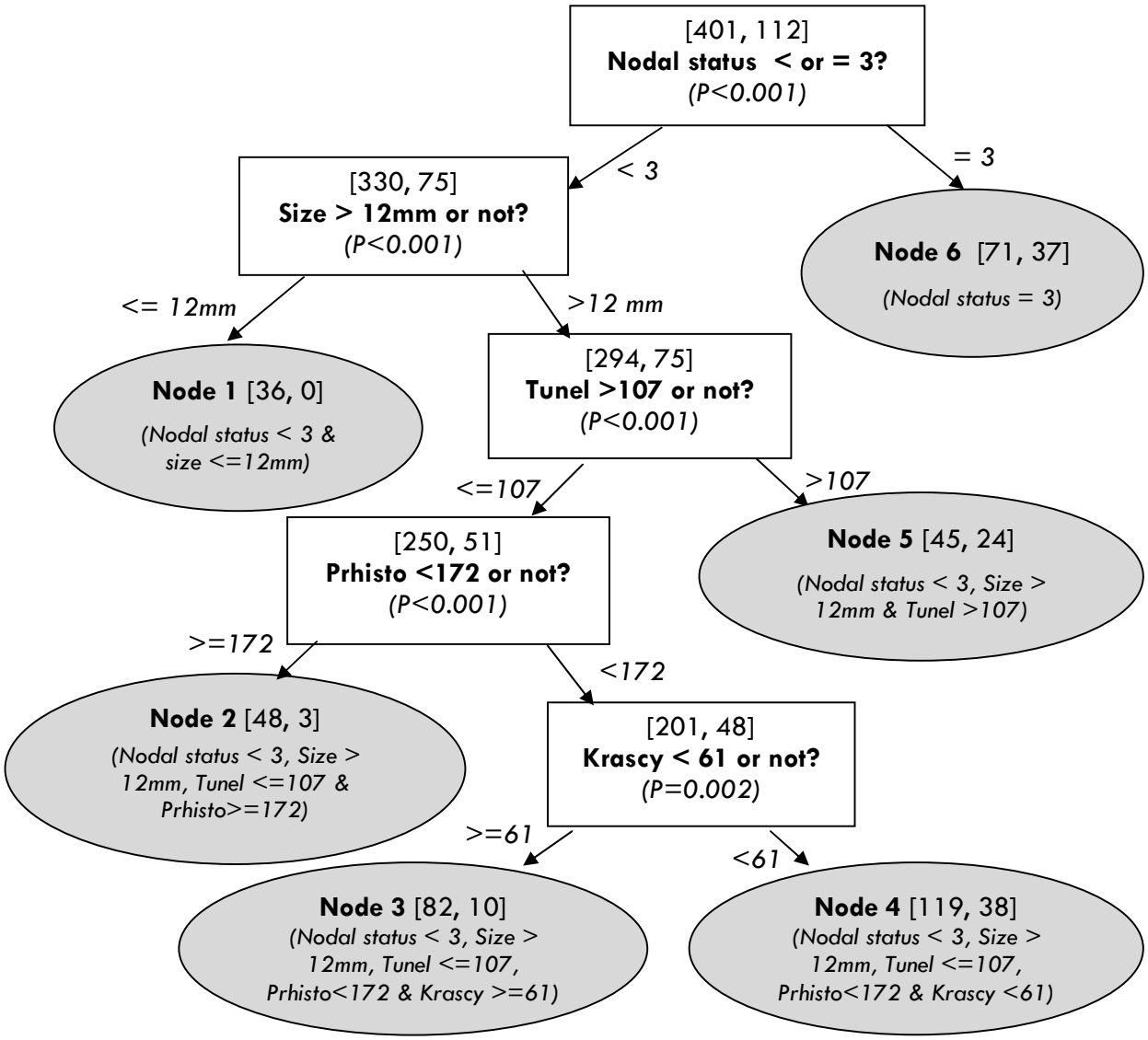

Figure 1. Classification tree using biomarkers and clinical predictors
For each branching point/ terminal node, numbers in square brackets show respectively number of patients and number of recurrences.
For each branching point, the Log-Rank P-value is reported for comparing RFS in resulting 'child' nodes.
For each terminal node, characteristics of cases within it are listed in curved brackets.

To improve the prognostic prediction, bagging of survival trees has been proposed. In this approach, a large number of trees are constructed by re-sampling from the original data. The aggregated Kaplan-Meier curve for a new patient is defined as the Kaplan-Meier curve of all observations identified by the M leaves containing the new patient [43]. Therefore, no single tree can be reported and communication of results is not simple. Application of bagging was beyond the scope of this paper. Another approach is to construct the tree using half of the data, and then investigate its ability in risk prediction in the second half (known as data-splitting technique). This approach is not practicable when sample size is small [44].

Missing data is a common problem in oncology. In a recent review of 100 papers reporting survival analysis, published in 2002, a total of 81 papers had data with missing covariates and 4 papers did not provide sufficient information to determine

Table 3. Estimated RFS rates in each of tree nodes

Node number	N at start	5-year RFS(%)	7-year RFS(%)	Risk grouping
1	36	100	100	Low
2	48	94	94	
3	82	93	87	Low Intermediate (LI)
4	119	74	67	High Intermediate (HI)
5	45	64	54	High
6	71	48	48	

whether there was missing data [45]. Complete-case analysis, which is the exclusion of cases with missing data on any of variables under study, was the method most frequently used [45]. However, exclusion of patients with missing data will diminish precision of estimates and genralisability of results [46]. The TSM approach to prognostic modelling has the advantage that the method handles missing data, and so intensive imputation techniques can be avoided.

Only 3.5% of patients in the lowest risk group experienced a recurrence compared to 53% of those categorised into the highest risk group had a recurrence (61 out of 116). Endocrine therapy, using tamoxifen or aromatase inhibitors remains the most successful approach to the treatment of early breast cancer but it is likely that many women do not require endocrine therapy at all, or if treated with aromatase inhibitors and/or chemotherapy, derive minimal additional benefit over tamoxifen treatment [47].

Molecular differences between breast cancer tumours support treating different molecular sub-types based on their biology and pathology rather than pathology alone. We have taken a purely statistical 'survival curve' approach to merging of small terminal nodes into larger ones. As can be seen from Figure 1 and Table 3, patients with large tumours and with low Tunel and High Prhisto had similar prognosis to that of patients with small tumours (which resulted in merging of nodes 1 and 2). While this might be clinically counter-intuitive, reflection on this could enable TSM models to offer insights, and generate new biological hypotheses about mechanisms that govern cancer progress and interactions between biomarkers, which could be tested in fresh samples. As another example, cases with nodal status of <3 and tumour size of <12 had a poor prognosis if their tunel is higher than 107, but had a very good prognosis if their tunel is lower than this threshold and if their Prhisto is higher than 172.

It is likely that the validity of the final refinements to the TSM model could be improved with cancer biologist input to decisions regarding merging of small terminal nodes (with similar recurrence-free survival curves), by taking into account the place of TMA variables along the cancer pathway. Certainly the most important issue for a model is its external validity, the extent to which it provides good predictions for similar patients who were not involved in the development of the model. If performance is assessed on the same sample as used for model development, then performance will be overestimated [48]. Therefore results presented in this paper are tentative until validated using a new data set.

Referring to the possibility that for some patients the balance of harm/benefit means that they would be better off with no adjuvant treatment, it has been postulated that 'it is an inability to identify such patients prior to treatment, rather than an expectation that all patients derive benefit, which drives the treatment of significant number of breast cancer patients with often aggressive chemotherapy' [49]. It has also been commented that the identification of novel prognostic markers and their integration in risk prediction is essential for the solution of this dilemma [49]. Over the past few years, applying regression modelling strategies, the role of a large number of candidate predictive biomarkers has been explored [14-16,50,51]. It has been concluded that tumour profiling might improve patient selection for endocrine therapies [14-16,50,51], and that over the next 3 to 5 years biomarkers will be incorporated as part of clinical diagnostic decision making [49]. The majority of published prognostic studies using TSM have focussed on clinical variables and traditional risk factors (such as race, ER and Progesterone Receptor (PR) status, family history) [24]. As an example of a model based on biomarkers, a total of 126 biomarkers, but no other variables, were available to construct a decision tree which predicts recurrence of breast

cancer [34]. The final tree used information on 6 of the biomarkers. However, to the best of our knowledge, our present study is the first to assess the value of integration of biological and clinical variables together into risk prediction, using TSM.

There are a variety of statistical approaches to deal with the issue of many potential variables [17], and elsewhere we have applied complex regression methods to develop a prognostic model combining biomarkers and clinical variables [41]. However, TSM analysis has a lot to offer researchers, because it provides a readily interpretable picture, which results in easier clinical decision making, and aids future studies [24,24,52-54]. For example, traditional multi-factorial regression tools (such as Cox without interaction term), suppose a uniform effect of the variable for the whole sample, whereas TSM can reveal a factor with different effects in different subgroups, a biologically plausible situation, in that a biomarker might be important for only a subset of patients. TSM therefore has potential benefits in terms of therapeutic management [24]. Furthermore, TSM is easy to apply (no distributional assumptions to be checked), avoids the needs for techniques to deal with missing data, and can be used as a good approximation for a complex model.

Acknowledgments

The data set analyzed was collected under the direction of Professor John Bartlett and his colleagues (Tove Kirkegaard, Liane McGlynn, and Timothy Cooke) at Glasgow University. The work was carried out as part of Mr Baneshi's PhD degree (submitted to Edinburgh University under supervision of Dr Pamela Warner and Dr Niall Anderson) for which he was supported by the Ministry of Health and Medical Education of Iran and Kerman Medical University.

Conflict of Interest

The authors declare no conflict of interest.

Authors' Contribution

The statistical ideas for the PhD study were BMR's. However his regular meetings with the supervisors (Dr WP and Dr AN) and the clinical collaborators of this study (in particular Professor BJ), to address the statistical approaches useful in the field of time-to-event clinical outcome data, and in presentation of results in a way meaningful to a clinical audience, will have contributed to the evolution of some ideas. All authors have read and approved the final version.

References

1. McPherson K, Steel CM, Dixon JM. ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics.BMJ 2000 Sep 9; 321(7261):624-8.

2. Cancer Research UK. Breast Cancer: UK breast cancer statistics. http://info cancerresearchuk org/cancerstats/types/breast/ 2006 November [cited 2007 Feb 26]; Available from: URL: http://info.cancerresearchuk.org/ cancerstats/ types/ breast/

3. Altman DG, Lyman GH. Methodological challenges in the evaluation of prognostic factors in breast cancer. Breast Cancer Res Treat 1998; 52(1-3):289-303.

4. Haybittle JL, Blamey RW, Elston CW, Johnson J, Doyle PJ, Campbell FC, et al. A prognostic index in primary breast cancer.Br J Cancer 1982 Mar; 45(3):361-6.

5. Todd JH, Dowle C, Williams MR, Elston CW, Ellis IO, Hinton CP, et al. Confirmation of a prognostic index in primary breast cancer. Br J Cancer 1987 Oct; 56(4):489-92.

6. Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham Prognostic Index in primary breast cancer. Breast Cancer Res Treat 1992; 22(3):207-19.

7. Balslev I, Axelsson CK, Zedeler K, Rasmussen BB, Carstensen B, Mouridsen HT. The Nottingham Prognostic Index applied to 9,149 patients from the studies of the Danish Breast Cancer Cooperative Group (DBCG). Breast Cancer Res Treat 1994; 32(3):281-90.

8. Lee AH, Ellis IO. The nottingham prognostic index for invasive carcinoma of the breast. Pathol Oncol Res 2008; 14(2):113-5.

9. Kirkegaard T, Bartlett JM. Novel pharmacodiagnostics in breast cancer. European oncological disease 2006; 53-6.

10. Martin M. Molecular biology of breast cancer. Clin Transl Oncol 2006 Jan; 8(1):7-14.

11. Chen W, Foran DJ. Advances in cancer tissue microarray technology: Towards improved understanding and diagnostics. Anal Chim Acta 2006 Mar 30; 564(1):74-81.

12. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol 1995 Dec; 48(12):1503-10.

13. Therneau TM, Atkinson EJ. An interoduction to recursive partitioning using the rpart routine. 1997 Feb 11.

14. McGlynn LM, Kirkegaard T, Edwards J, Tovey S, Cameron D, Twelves C, et al. Ras/Raf-1/MAPK pathway mediates response to tamoxifen but not chemotherapy in breast cancer patients. Clin Cancer Res 2009 Feb 15; 15(4):1487-95.

15. Kirkegaard T, McGlynn LM, Campbell FM, Muller S, Tovey SM, Dunne B, et al. Amplified in breast cancer 1 in human epidermal growth factor receptor - positive tumors of tamoxifen-treated breast cancer patients. Clin Cancer Res 2007 Mar 1; 13(5):1405-11.

16. Kirkegaard T, Witton CJ, McGlynn LM, Tovey SM, Dunne B, Lyon A, et al. AKT activation predicts outcome in breast cancer patients treated with tamoxifen. J Pathol 2005 Oct; 207(2):139-46.

17. Baneshi MR. Statistical Models in Prognostic Modelling of Many Skewed Variables and Missing Data: A Case Study in Breast Cancer (PhD thesis submitted at Edinburgh University) 2009.

18. Williams BA, Mandrekar JN, Mandrekar SJ, Cha SS, Furth AF. Finding optimal cutpoints for continuous covariates with binary and time-to-event outcomes. 2006.

19. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part IV: further concepts and methods in survival analysis.Br J Cancer 2003 Sep 1; 89(5):781-6.

20. Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. J Natl Cancer Inst 1994 Jun 1; 86(11):829-35.

21. Radespiel-Troger M, Rabenstein T, Schneider HT, Lausen B. Comparison of tree-based methods for prognostic stratification of survival data. Artif Intell Med 2003 Jul; 28(3):323-41.

22. Dannegger F. Tree stability diagnostics and some remedies for instability. Stat Med 2000 Feb 29; 19(4):475-91.

23. Segal MR, Bloch DA. A comparison of estimated proportional hazards models and regression trees. Stat Med 1989 May; 8(5):539-50.

24. Banerjee M, George J, Song EY, Roy A, Hryniuk W. Tree-based model for breast cancer prognostication. J Clin Oncol 2004 Jul 1; 22(13):2567-75.

25. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. Br J Cancer 1957 Sep; 11(3):359-77.

26. Van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis.Stat Med 1999 Mar 30; 18(6):681-94.

27. Van Buuren S, Oudshoorn K. Multiple imputation by chained equations:MICE V 1.0 User's manual. 2000.

28. Blamey RW, Ellis IO, Pinder SE, Lee AH, Macmillan RD, Morgan DA, et al. Survival of invasive breast cancer according to the Nottingham Prognostic Index in cases diagnosed in 1990-1999. Eur J Cancer 2007 Jul; 43(10):1548-55.

29. Blamey RW, Pinder SE, Ball GR, Ellis IO, Elston CW, Mitchell MJ, et al. Reading the prognosis of the individual with breast cancer.Eur J Cancer 2007 Jul; 43(10):1545-7.

30. Sauerbrei W, Hubner K, Schmoor C, Schumacher M. Validation of existing and development of new prognostic classification schemes in node negative breast cancer. German Breast Cancer Study Group. Breast Cancer Res Treat 1997 Jan; 42(2):149-63.

31. Coradini D, Boracchi P, Daidone MG, Pellizzaro C, Miodini P, Ammatuna M, et al. Contribution of vascular endothelial growth factor to the Nottingham prognostic index in node-negative breast cancer. Br J Cancer 2001 Sep 14; 85(6):795-7.

32. Okugawa H, Yamamoto D, Uemura Y, Sakaida N, Yamada M, Tanaka K, et al. Prognostic factors in breast cancer: the value of the Nottingham Prognostic Index for patients treated in a single institution. Surg Today 2005; 35(11):907-11.

33. Kollias J, Murphy CA, Elston CW, Ellis IO, Robertson JF, Blamey RW. The prognosis of small primary breast cancers. Eur J Cancer 1999 Jun; 35(6):908-12.

34. Ring BZ, Seitz RS, Beck R, Shasteen WJ, Tarr SM, Cheang MC, et al. Novel prognostic immunohistochemical biomarker panel for estrogen receptor-positive breast cancer.J Clin Oncol 2006 Jul 1; 24(19):3039-47.

35. Lundin J, Lehtimaki T, Lundin M, Holli K, Elomaa L, Turpeenniemi-Hujanen T, et al. Generalisability of survival estimates for patients with breast cancer--a comparison across two population-based series. Eur J Cancer 2006 Dec; 42(18):3228-35.

36. Sidoni A, Bellezza G, Cavaliere A, Del SR, Scheibel M, Bucciarelli E. Prognostic indexes in breast cancer: comparison of the Nottingham and Adelaide indexes. Breast 2004 Feb; 13(1):23-7.

37. Eden P, Ritz C, Rose C, Ferno M, Peterson C. "Good Old" clinical markers have similar power in breast cancer prognosis as microarray gene expression profilers. Eur J Cancer 2004 Aug; 40(12):1837-41.

38. D'Eredita' G, Giardina C, Martellotta M, Natale T, Ferrarese F. Prognostic factors in breast cancer: the predictive value of the Nottingham Prognostic Index in patients with a long-term follow-up that were treated in a single institution. Eur J Cancer 2001 Mar;37(5):591-6.

39. Brown J, Jones M, Benson EA. Comment on the Nottingham Prognostic Index. Breast Cancer Res Treat 1993; 25(3):283.

40. Callagy GM, Pharoah PD, Pinder SE, Hsu FD, Nielsen TO, Ragaz J, et al. Bcl-2 is a prognostic marker in breast cancer independently of the Nottingham Prognostic Index.Clin Cancer Res 2006 Apr 15; 12(8):2468-75.

41. Baneshi MR, Warner P, Anderson N, Bartlett JSM. Tamoxifen resistance in early breast cancer: statistical modelling of tissue markers to improve risk prediction. Br J Cancer 2010; 102: 1503-10.

42. Lausen B, Horton T, Bertz F, Schumacher M. Assessment of optimal selected prognostic factors. Biometrical Journal 2004; 46:364-74.

43. Hothorn T, Lausen B, Benner A, Radespiel-Troger M. Bagging survival trees. Stat Med 2004 Jan 15;23(1):77-91.

44. Steyerberg EW, Harrell FE, Jr., Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001 Aug; 54(8):774-81.

45. Burton A, Altman DG. Missing covariate data within cancer prognostic studies: a review of current reporting and proposed guidelines. Br J Cancer 2004 Jul 5; 91(1):4-8.

46. Kristman VL, Manno M, Cote P. Methods to account for attrition in longitudinal data: do they work? A simulation study. Eur J Epidemiol 2005; 20(8):657-62.

47. Abe O, Abe R, Enomoto K, Kikuchi K, Koyama H, Masuda H, et al. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-

year survival: an overview of the randomised trials. Lancet 2005 May 14; 365(9472):1687-717.

48. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med 1999 Mar 16; 130(6):515-24.

49. Faratian D, Bartlett JM. Predictive markers in breast cancer--the future.Histopathology 2008 Jan; 52(1):91-8.

50. Cannings E, Kirkegaard T, Tovey SM, Dunne B, Cooke TG, Bartlett JM. Bad expression predicts outcome in patients treated with tamoxifen. Breast Cancer Res Treat 2007 Apr; 102(2):173-9.

51. Tovey SM, Dunne B, Witton CJ, Forsyth A, Cooke TG, Bartlett JM. Can molecular markers predict when to implement treatment with aromatase inhibitors in invasive breast cancer? Clin Cancer Res 2005 Jul 1; 11(13):4835-42.

52. Harrell FE, Margolis PA, Gove S, Mason KE, Mulholland EK, Lehmann D, et al. Development of a clinical prediction model for an ordinal outcome: the World Health Organization Multicentre Study of Clinical Signs and Etiological agents of Pneumonia, Sepsis and Meningitis in Young Infants. WHO/ARI Young Infant Multicentre Study Group. Stat Med 1998 Apr 30; 17(8):909-44.

53. Ciampi A, Lawless JF, McKinney SM, Singhal K. Regression and recursive partition strategies in the analysis of medical survival data. J Clin Epidemiol 1988; 41(8):737-48.

54. Ture M, Tokatli F, Kurt I. Using Kaplan-Meier analysis together with decision tree methods (C&RT, CHAID, QUEST, C4.5 AND ID3) in determining recurrence free survival of breast cancer patients. Expert Systems with Applications 2009; 36:2017-26.