View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Tilburg University Repository

Tilburg University

Relationship between age and axillary lymph node involvement in women with breast cancer

Wildiers, H.; van Calster, B.; van de Poll-Franse, L.V.; Hendrikx, W.; Røislien, J.; Smeets, A.; Paridaens, R.; Deraedt, K.; Leunen, K.; Weltens, C.; Van Huffel, S.; Christiaens, M.R.; Neven, P.

Published in:
Journal of Clinical Oncology

Publication date:
2009

Document Version
Publisher's PDF, also known as Version of record

Link to publication in Tilburg University Research Portal

Citation for published version (APA):
Wildiers, H., van Calster, B., van de Poll-Franse, L. V., Hendrikx, W., Røislien, J., Smeets, A., Paridaens, R., Deraedt, K., Leunen, K., Weltens, C., Van Huffel, S., Christiaens, M. R., & Neven, P. (2009). Relationship between age and axillary lymph node involvement in women with breast cancer. *Journal of Clinical Oncology*, *27*(18), 2931-2937.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 12. May. 2021

Published Ahead of Print on May 18, 2009 as 10.1200/JCO.2008.16.7619
The latest version is at http://jco.ascopubs.org/cgi/doi/10.1200/JCO.2008.16.7619

JOURNAL OF CLINICAL ONCOLOGY ORIGINAL REPORT

Relationship Between Age and Axillary Lymph Node Involvement in Women With Breast Cancer

Hans Wildiers, Ben Van Calster, Lonneke V. van de Poll-Franse, Wouter Hendrickx, Jo Røislien, Ann Smeets, Robert Paridaens, Karen Deraedt, Karin Leunen, Caroline Weltens, Sabine Van Huffel, Marie-Rose Christiaens, and Patrick Neven

From the Multidisciplinary Breast Centre and Departments of General Medical Oncology, Pathology, Gynecology, and Radiotherapy, University Hospitals Leuven; Department of Electrical Engineering, Katholieke Universiteit Leuven, Leuven, Belgium; Comprehensive Cancer Centre South, Eindhoven Cancer Registry, Eindhoven, the Netherlands; and Department of Biostatistics, Institute of Basic Medical Sciences, University of Oslo, Norway.

Submitted February 13, 2008; accepted January 29, 2009; published online ahead of print at www.jco.org on May 18, 2009.

Supported by Katholieke Universiteit Leuven research council Grant No. GOA-AMBioRICS and European Union projects BIOPATTERN (Grant No. FP6-2002-IST 508803), ETUMOR (Grant No. FP6-2002-LIFESCIHEALTH 503094), and Healthagents (Grant No. IST–2004-27214). B.V.C. is a postdoctoral researcher funded by the Research Foundation–Flanders (FWO).

H.W. and B.V.C both contributed equally to this work.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Hans Wildiers, MD, PhD, University Hospitals Leuven, Herestraat 49, Leuven B-3000, Belgium; e-mail: hans.wildiers@uz.kuleuven.ac.be.

© 2009 by American Society of Clinical Oncology

0732-183X/09/2799-1/$20.00

DOI: 10.1200/JCO.2008.16.7619

ABSTRACT

Purpose
To study the relation between the presence of axillary lymph node (LN) involvement and age in breast cancer.

Patients and Methods
The breast cancer database of the University Hospitals Leuven contains complete data on 2,227 patients with early breast cancer consecutively treated between 2000 and 2005. A multivariate piecewise logistic regression model was used to analyze LN involvement in relation to age at diagnosis. A similar analysis was then performed on a large, independent, population-based database from the Eindhoven Cancer Registry to investigate whether the effects of the Leuven model could be replicated.

Results
We observed a piecewise effect of age. That is, women up to 70 years of age were less likely to have positive LNs with increasing age (odds ratio per 10-year increase, 0.87). In contrast, older women were more likely to have positive LNs with increasing age. However, for older women, the effect of age interacted with tumor size ($P = .0044$), suggesting that increasing age is associated with increased risk of LN involvement, mainly in small tumors. These findings were replicated in the Eindhoven Cancer Registry database.

Conclusion
Axillary LN involvement varies with age at diagnosis; its probability decreases with increasing age up to the age of approximately 70 years, but increases again thereafter. However, this increase is mainly seen in smaller tumors and suggests a different behavior of small breast cancers in older adult patients. We hypothesize that decreased immune defense mechanisms, related with aging, may play a role in earlier invasion into LNs.

J Clin Oncol 27. © 2009 by American Society of Clinical Oncology

INTRODUCTION

Breast cancer is the most frequently occurring tumor in older women in Europe and North America.[1] Advancing age is associated with more favorable tumor biology indicated by increased hormone sensitivity, less HER-2/*neu* overexpression, lower grading, and lower proliferation indices.[2-4] However, at the time of diagnosis, older adult patients more often have an advanced-stage breast cancer and larger tumors.[1]

There has been conflicting data on lymph node involvement and aging; some studies showed decreased involvement in older adult patients,[5-7] whereas others did not find an effect[3,4,8] or showed increased involvement with increasing age.[1,9] Because of discordant results between previous studies, we decided to perform a retrospective analysis assessing the connection between age and lymph node involvement, while adjusting for other clinicopathologic variables.

PATIENTS AND METHODS

Patients

The breast cancer database from the University Hospitals (UH) Leuven in Leuven, Belgium, contains data on 3,549 patients from the period between 2000 and 2005. Patients for the present retrospective study were first selected according to the following inclusion criteria: female patients diagnosed with invasive breast cancer; patients who had received primary surgery at the UH Leuven (thus excluding patients who had neoadjuvant systemic therapy and patients who never had therapy/surgery before for breast cancer); patients whose pathology was performed in the UH Leuven; patients with no metastases at diagnosis. In total, 2,568 patients fulfilled these criteria.

Information downloaded from jco.ascopubs.org and provided by CATHARINA ZIEKENHUIS on June 3, 2009 from 137.56.139.73.
Copyright © 2009 by the American Society of Clinical Oncology. All rights reserved.
Copyright 2009 by American Society of Clinical Oncology

Variables Studied

The following variables were included for analysis: age at diagnosis, maximal microscopic tumor size (largest diameter), worst tumor grade, axillary lymph node status including number of positive lymph nodes, estrogen receptor (ER) and progesterone receptor (PR) status, and HER-2/*neu* status. Lymph node status was analyzed as a binary variable indicating whether positive lymph nodes were detected (ie, lymph node involvement). Determination of tumor grading and ER, PR, and HER-2/*neu* status was done according to established procedures, which are described in more detail elsewhere.[10]

There were missing values for tumor size (n = 97; 3.8%), tumor grade (n = 16; 0.6%), receptor status (n = 146; 5.7%), and axillary lymph node status (n = 82; 3.2%). The 2,227 patients (86.7%) with complete information were used for statistical analysis. Of the 82 patients with missing lymph node status, 48 patients did not receive an axillary lymph node dissection (pNx); for the remaining 34 patients, information on axillary lymph node dissection was missing.

Statistical Analysis of the Leuven Data

First, descriptive statistics were computed, and univariate relationships between all aforementioned variables and lymph node involvement were evaluated using logistic regression. Logistic regression results are presented as odds ratios (OR) with 95% CIs, unless stated otherwise. Because a nonlinear relationship between age and lymph node involvement was expected based on existing literature, lymph node involvement was also regressed on age using nonparametric logistic regression based on locally weighted scatterplot smoothing (lowess).[11] This allowed us to investigate the true functional form of the relationship between age and lymph node status.

Thereafter, multivariate logistic regression was used to predict lymph node involvement. The following variables were considered as predictors in the multivariate model: age at diagnosis (years), tumor size (millimeters), tumor grade (1, 2, or 3), and combined ER, PR, and HER-2/*neu* receptor status (six categoric levels: ER-negative/PR-negative /HER-2–negative, ER-negative/PR-negative/HER-2–positive, ER-positive/PR-negative/HER-2–negative, ER-positive/PR-negative/HER-2–positive, ER-positive/PR-positive/HER-2–negative, ER-positive/PR-positive/HER-2–positive). Details on model and variable selection are provided in the Appendix (online only).

Replication of the Model Using the Eindhoven Cancer Registry Database

The logistic regression model predicting lymph node involvement based on the Leuven data seemed both robust and well-fitting (see Results). We therefore investigated whether the estimated effects of age and tumor size would be replicated on an independent tumor database from the Eindhoven Cancer Registry (ECR), Eindhoven, the Netherlands. Since 1955, the ECR records data on all patients newly diagnosed with cancer in the southern part of the Netherlands (covering an area with 2.3 million inhabitants, 17 hospital locations, and two large radiotherapy institutes; there are no university hospitals in the area). The ECR routinely collects data on tumor characteristics such as date of diagnosis, subsite, histology, stage (TNM clinical classification), primary treatment, and patient characteristics (eg, sex and date of birth).

A subset of the full data set, containing 3,234 patients diagnosed with invasive breast cancer in 2005 and 2006, was used for the analysis. These patients fulfilled the inclusion criteria applied to the Leuven database. We chose not to use data from the period before 2005 since variables were recorded more systematically since the end of 2004. Data on lymph node involvement, ER, PR, and HER-2 status, tumor grade, and tumor size was missing in 9.2%, 2.9%, 3.8%, 18.5%, 16.4%, and 12.5% of the patients, respectively. As a result, complete information was available for only 2,155 patients (66.6%). Because of the large amount of missing values, which increases the risk for biased results, we imputed the missing values using multiple imputation (MI).[13,14] MI imputes missing values multiple times to account for uncertainty in the imputations. Each imputation of the missing values results in a complete ECR data set. We chose to impute missing values five times, as this is usually sufficient.[12,13] We thus had five complete ECR data sets on which to perform the statistical analysis. Following standard MI practice,[12,13] all analyses were performed on each complete ECR data set, and the results were combined to obtain the final results. See Appendix (online only) for details on the imputation procedure.

RESULTS

The tumor characteristics of the patients from UH Leuven and ECR are shown in Tables 1 and 2. Here, age was arbitrarily divided into four groups: younger than 50 years, 50 to 69 years, 70 to 79 years, and 80 years or older.

Univariate Analysis

Age had a fairly weak relationship with lymph node involvement (OR per 10-year increase 0.91; 95% CI, 0.85 to 0.98). However, the univariate nonparametric logistic regression suggested a piecewise effect of age on lymph node status (Fig 1A). Age clearly seems to be negatively related to lymph node involvement for women up to approximately 70 years and positively related for women who are older than 70 years at diagnosis. Tumor size (OR per cm increase 1.58; 95% CI, 1.49 to 1.68), tumor grade (OR per grade increase 1.64; 95% CI, 1.44 to 1.87), and HER-2/*neu* receptor status (OR, 1.51; 95% CI, 1.16 to 1.96) had clear positive relationships with lymph node involvement. ER (OR 1.09; 95% CI, 0.84 to 1.42) and PR receptor status (OR, 1.14; 95% CI, 0.93 to 1.40) were not clearly related to the outcome.

Multivariate Logistic Regression

The nonlinear relationship between age and lymph node involvement was still present when adjusting for the other predictors. Once again, age had a piecewise effect on lymph node involvement, with the effect reversing at approximately the age of 70 years. This particular cutoff is also frequently used in clinical studies to separate older adult from younger patients and has been suggested as a cutoff for geriatric assessment.[14] Thus it was decided to use a piecewise logistic regression model[15] with a different effect of age for patients up to 70 years and patients older than 70 years. It is important to emphasize that we did not fit two separate models, but rather one single piecewise logistic regression model where we modeled age as changing its effect on lymph node involvement at 70 years.

Standard statistical variable selection criteria, such as Akaike's information criterion and Bayesian information criterion[16] and likelihood ratio *P* values (see Appendix, online only, for details), suggested that the optimal model was obtained when keeping all predictors in the model. However, ER-positive/PR-positive/HER-2 –positive tumors seemed more often spread to the lymph nodes than other tumors. A likelihood ratio χ^2 test suggested that the model could indeed be simplified by reducing the combined ER, PR, and HER-2/*neu* receptor status to a binary variable indicating whether all three receptor statuses were positive or not (triple positivity). Next, the necessity of interaction terms was investigated (see Appendix, online only, for details). Interactions involving age were investigated before and after the breakpoint separately. One interaction term was included in the final model: for women older than 70 years, the effect of age seems to interact with tumor size ($P = .0044$). The final model is presented in Table 2. Model fit was good, as assessed by the Hosmer-Lemeshow goodness-of-fit test. See Appendix, online only, for the model equation.

Up to 70 years of age, the odds of lymph node involvement decreased by 13% for each 10-year increase in age (OR, 0.87; 95% CI,

 © 2009 by American Society of Clinical Oncology

Information downloaded from jco.ascopubs.org and provided by CATHARINA ZIEKENHUIS on June 3, 2009 from 137.56.139.73.
Copyright © 2009 by the American Society of Clinical Oncology. All rights reserved.

Table 1. Tumor Characteristics and Lymph Node Involvement in Relation to Age

Parameter	Age Category (years)				
	< 50	50-69	70-79	≥ 80	Total
University Hospitals Leuven, No.	636	1,189	310	92	2,227
Histology, %					
Ductal	86.8	84.7	82.0	82.0	84.9
Lobular	12.0	14.5	15.4	18.0	14.0
Other	1.2	0.8	2.6	0	1.1
Median tumor size, mm	22	20	24.5	30	20
Tumor grading, %					
1	9.7	18.2	10.6	7.6	14.3
2	41.2	47.3	49.7	50.0	46.0
3	49.1	34.5	39.7	42.4	39.7
ER+, %	85.1	88.6	87.1	88.0	87.4
PR+, %	79.4	76.8	74.2	73.9	77.1
HER-2+, %	14.3	10.9	8.4	9.8	11.5
Lymph node involvement, %	40.5	35.4	32.2	43.2	37.0
No positive lymph nodes	59.5	64.6	64.8	56.8	62.8
1-3 positive lymph nodes	26.7	25.3	23.2	28.4	25.6
≥ 4 positive lymph nodes	13.8	10.1	12.0	14.8	11.6
Eindhoven Cancer Registry, No.	831	1,535	508	360	3,234
Histology, %					
Ductal	83.8	80.5	77.8	81.4	81.0
Lobular	8.3	10.9	14.8	11.1	10.9
Other	7.9	8.6	7.5	7.5	8.1
Median tumor size, mm	19	16	17	23	18
Tumor grading, %					
1	21.3	30.3	25.7	28.7	27.1
2	39.8	45.0	52.8	48.2	45.2
3	38.9	24.7	21.5	23.2	27.7
ER+, %	77.9	83.1	90.4	88.9	83.6
PR+, %	71.7	68.5	76.1	76.1	71.3
HER-2+, %	18.7	15.2	11.0	10.8	15.0
Lymph node involvement, %	48.2	38.2	36.5	44.6	41.2
No positive lymph nodes	51.8	61.8	63.5	55.4	58.8
1-3 positive lymph nodes	35.2	26.6	27.0	34.1	29.7
≥ 4 positive lymph nodes	13.0	11.6	9.5	10.5	11.5

Abbreviations: ER, estrogen receptor; PR, progesterone receptor.

0.78 to 0.96) and increased by 60% for each centimeter increase in tumor size (OR, 1.60; 95% CI, 1.50 to 1.70). After the age of 70 years, increasing age increased the risk of lymph node involvement, as did increasing tumor size. However, for older women, the effects of age and tumor size were not independent because of the interaction term. This interaction term can be explained from two different angles. Focusing on the effect of age, the positive effect of age in older women was most pronounced for small tumors. On the contrary, for very

Table 2. Multivariate Logistic Regression Model for Lymph Node Involvement, With Piecewise Effect of Age and Interaction Between Age and Tumor Size for Women Older Than 70 Years

Variable	Leuven University Hospital (n = 2,227)			Eindhoven Cancer Registry (n = 3,234)		
	OR	95% CI	*P*	OR	95% CI	*P*
Age, years (for women ≤ 70)	0.87*	0.78 to 0.96	.0048	0.82*	0.76 to 0.90	< .0001
Age, years (for women older than 70)	2.93*	1.60 to 5.35	.0008	1.84*	1.25 to 2.71	.0023
Size, mm	1.60*	1.50 to 1.70	< .0001	1.89*	1.69 to 2.11	< .0001
Grade	1.28	1.10 to 1.48	.0009	1.19	1.03 to 1.36	.0172
Triple positive	2.09	1.43 to 3.06	.0001	1.06	0.76 to 1.46	.7373
Age · size (for women older than 70 years)	0.78*	0.66 to 0.92	.0044	0.81*	0.67 to 0.99	.0414

NOTE. Model development on data from University Hospitals Leuven; model replication on data from the Eindhoven Cancer Registry.
Abbreviation: OR, odds ratio.
*Expressed per 10-year increase in age and/or per centimeter increase in size, as applicable.

© 2009 by American Society of Clinical Oncology

Information downloaded from jco.ascopubs.org and provided by CATHARINA ZIEKENHUIS on June 3, 2009 from 137.56.139.73.
Copyright © 2009 by the American Society of Clinical Oncology. All rights reserved.

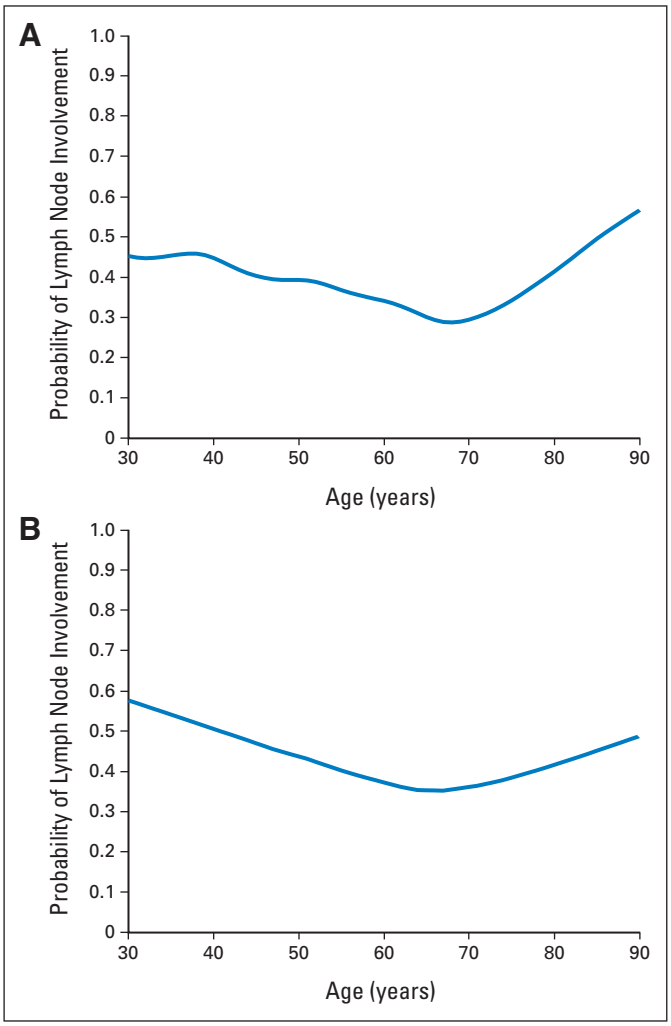

Fig 1. Relation between age and lymph node involvement based on univariate nonparametric logistic regression: (A) University Hospital Leuven data (n = 2,227) and (B) Eindhoven Cancer Registry data (n = 3234).

large tumors, the risk of lymph node involvement even decreased with increasing age. Focusing on the effect of tumor size, the positive effect of tumor size in older women attenuated when age at diagnosis increased. This joint effect of age and tumor size for older women is demonstrated in Figure 2A, which shows the predicted probability of lymph node involvement as a function of age for three arbitrarily chosen groups according to tumor size: patients with a tumor size up to 20 mm (compatible with T1 according to TNM staging system), patients with a tumor size between 21 and 35 mm, and patients with a tumor size larger than 35 mm. Finally, tumor grade and triple positivity each had an independent and general positive effect on lymph node involvement (Table 2).

Tumor characteristics and lymph node involvement in relation to tumor size are shown in Table 3. Here, tumor size was arbitrarily divided into the same three groups as previously mentioned: $\leq$ 20 mm, 21 to 35 mm, and more than 35 mm. In general, lymph node involvement was much more frequent in larger tumors (22.0%, 40.9%, and 67.4%, respectively). When looking in more detail, this tendency was stronger for younger than for older patients. Also, for large tumors, nodal involvement decreased with increasing age at diagnosis (eg, 53.3% for $\geq$ 80-year-old patients with a tumor larger than 35 mm *v* 75.2% for patients $<$ 50 years of age). For smaller tumors, a different picture was seen: nodal involvement was less frequent in younger compared with older patients.

Replication of the Effects of Age and Tumor Size on the ECR Database

The univariate effect of age on lymph node involvement using nonparametric regression is shown in Figure 1B. In the ECR data, a similar piecewise effect of age on lymph node involvement was observed, but possibly with a slightly lower breakpoint. Then, a multivariate logistic regression model with the same predictors was fitted to investigate whether the effects of age and tumor size could be replicated on an independent breast cancer database. The piecewise effect of age and the interaction between age and tumor size for older women were replicated (Table 2 and Fig 2B), even though the interaction effect is slightly weaker compared with the Leuven model (see Appendix, online only, for the model equation). Tumor characteristics and lymph node involvement in relation to tumor size are shown in Table 3.

DISCUSSION

This study investigated the relation between lymph node involvement and age in a large database from a single center (UH Leuven). Univariate analysis showed a slight decrease in lymph node involvement with age when age was considered as having a linear effect throughout its whole range. However, nonparametric analysis suggested a piecewise effect of age with a breakpoint of approximately 70 years. In multivariate logistic regression, modeling the effect of age as being piecewise linear, higher age at diagnosis decreased the risk of lymph node involvement up to the age of 70 years, but increased the risk after the age of 70 years. The effect of age after 70 years, however, interacted with tumor size. More specifically, from the point of view of age at diagnosis, the positive effect of age was most pronounced in small tumors. For large tumors, increasing age even decreased the risk of lymph node involvement. From the point of view of tumor size, the positive effect of this predictor on lymph node involvement became weaker as age at diagnosis increased.

These findings were replicated in an independent data set of 3,234 women registered with breast cancer in the ECR. This successful replication is a major strength of this study. Together, these results provide good evidence that there is a decrease in lymph node involvement up to a certain age, with an increase thereafter (mainly in small tumors). The effect of age on lymph node involvement by age is nonlinear.

An explanation for this nonlinearity is not evident, and it is possible that two or more influencing factors, acting in a different way, are responsible for the effect. One explanation could be that breast tumors are detected in a later stage in older adult patients. There is no systematic breast cancer screening at ages older than 70 years in Belgium. Moreover, older adult patients wait longer before consulting a physician.[17] Consistent with existing literature,[3,9] we observed that tumors were larger at diagnosis in older adults. Because there is a clear and well-known positive correlation between tumor size and lymph

 © 2009 by American Society of Clinical Oncology

Information downloaded from jco.ascopubs.org and provided by CATHARINA ZIEKENHUIS on June 3, 2009 from 137.56.139.73.
Copyright © 2009 by the American Society of Clinical Oncology. All rights reserved.

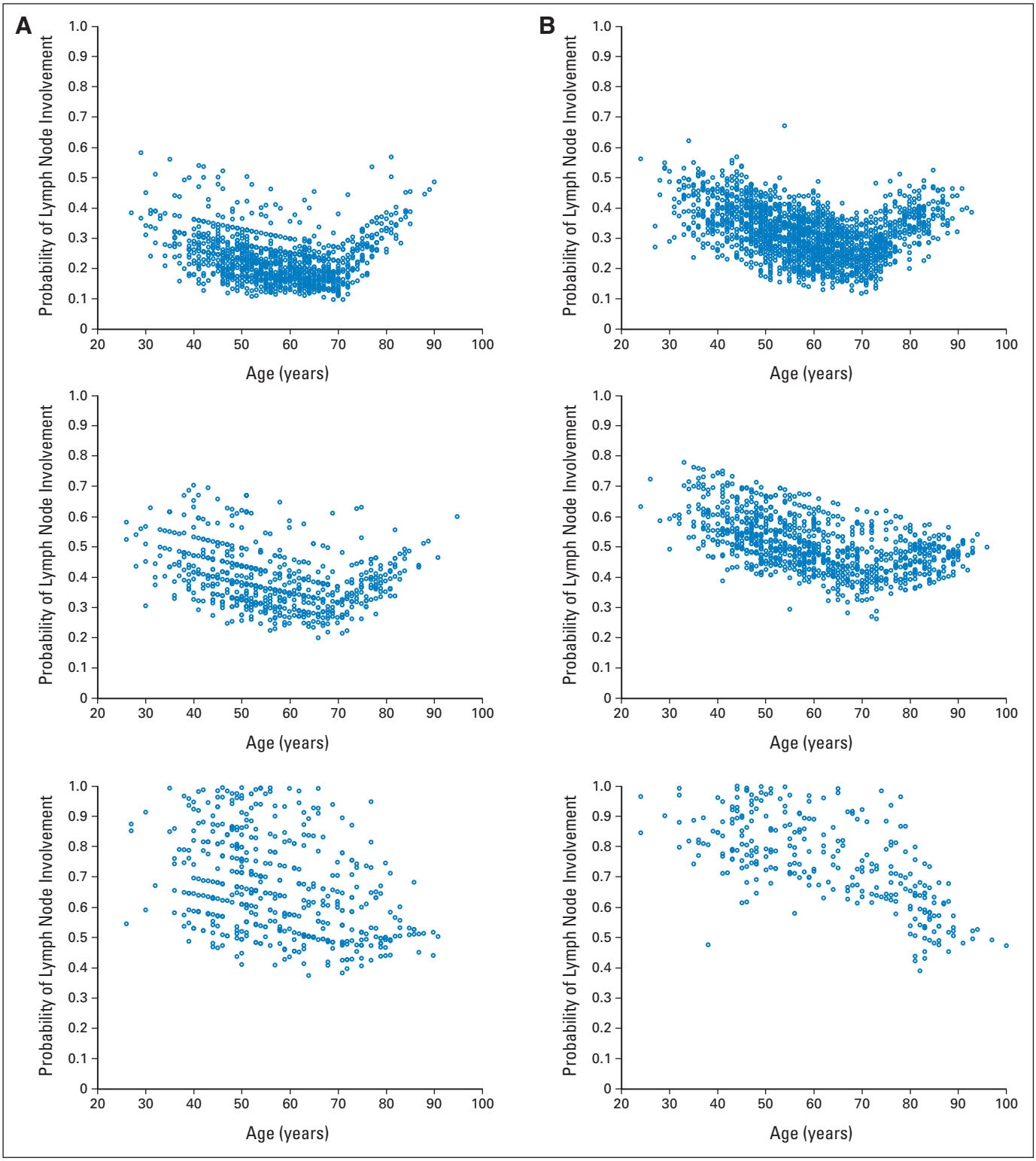

Fig 2. Scatterplots between age and the model's predicted probability of lymph node involvement. Separate plots are shown for (A) the University Hospital Leuven data (n = 2,227, left plots) and (B) Eindhoven Cancer Registry data (n = 3,234, right plots), and for women with tumor size up to 20 mm (top plots), tumor size between 21 and 35 mm (middle plots), and tumor size greater than 35 mm (bottom plots).

node involvement, the risk of lymph node involvement would thus increase in older adults. However, there was an interaction between age and tumor size on lymph node involvement for older women, suggesting that the effect of age on lymph node involvement is not independent of tumor size. More specifically, the interaction suggested that the positive effect of age for older women is most pronounced in smaller tumors ($\leq$ 20 mm), whereas for large tumors, the effect of age even became negative. This refutes the hypothesis

© 2009 by American Society of Clinical Oncology

Information downloaded from jco.ascopubs.org and provided by CATHARINA ZIEKENHUIS on June 3, 2009 from 137.56.139.73.
Copyright © 2009 by the American Society of Clinical Oncology. All rights reserved.

Table 3. Tumor Characteristics and Lymph Node Involvement in Relation to Tumor Size and Age

Parameter	Tumor Size (mm)					
	≤ 20		21-35		> 35	
	No.	%	No.	%	No.	%
University Hospitals Leuven, No.	1,115		646		466	
Median age, years	57		57		56	
Tumor grade						
1		22.0		9.1		3.2
2		48.2		40.3		48.7
3		29.9		50.6		48.1
ER+		88.5		85.9		86.7
PR+		78.1		75.1		77.3
HER-2+		10.0		13.5		12.2
Lymph node involvement per age group						
< 50 years	296	22.0	195	45.6	145	75.2
50-69 years	654	21.3	317	39.4	218	67.4
70-79 years	134	22.4	103	33.0	73	57.5
80+ years	31	35.5	31	51.6	30	53.3
Overall		22.0		40.9		67.4
Eindhoven Cancer Registry	1,936		1,014		284	
Median age, years	59		58		62	
Tumor grade						
1		33.5		18.8		12.7
2		46.6		43.0		43.9
3		19.9		38.2		43.4
ER+		87.6		79.1		72.3
PR+		74.8		67.7		60.3
HER-2+		13.0		17.9		17.5
Lymph node involvement per age group, years						
< 50	460	38.3	291	54.9	80	80.9
50-69	1,015	27.4	432	55.7	88	77.4
70-79	315	24.7	145	50.9	48	71.1
80+	146	31.9	146	50.5	68	59.5
Overall		29.9		54.0		73.0

Abbreviations: ER, estrogen receptor; PR, progesterone receptor.

that the increased lymph node involvement in older adults is only related to the presence of larger tumors. It seems that it is mainly the smaller tumors in older adult patients that metastasize more frequently to the lymph nodes than in younger postmenopausal patients, whereas the larger tumors in older adult patients metastasize less frequently. Or, differently stated, if breast tumors in older adults have the capacity to metastasize to lymph nodes, this occurs in an earlier stage/smaller tumors than in younger postmenopausal patients.

One might consider selection bias as a cause of the age-related effects. Axillary surgery is sometimes omitted in older adult patients with small tumors and a clinically node-negative axilla, because previous studies have not been able to show survival benefit from axillary surgery in this population.[18,19] This could lead to a higher percentage of lymph node involvement in the older adult population where an axillary dissection was performed. However, an axillary dissection was not performed in only 48 patients because of several reasons (eg, small unexpected invasive foci, unfit for general anesthesia), and 26 of these 48 patients were ≤ 70 years of age. Selection bias, if present, was thus probably limited. This is supported by the observation that the effects concerning age and size were confirmed in an independent database where missing values were imputed using multiple imputation.

In the Leuven data, the odds of lymph node involvement for triple positive tumors (ER, PR, HER-2 positive) was 2.09 times higher than for other tumors (95% CI, 1.43 to 3.06). In the ECR data, however, the OR was only 1.06 (95% CI, 0.76 to 1.46), suggesting a lack of effect. We do not have a full explanation for this difference, but major differences in immunohistochemical techniques and cutoff values for positivity throughout the 17 different Dutch hospital settings might have had serious impact on these results.

The piecewise effect for age required the specification of a breakpoint. However, in reality, we are probably not dealing with a breakpoint in the hard sense but rather with a transition phase that was best modeled using a piecewise effect. A breakpoint of 70 years was used in this study because the turning point was at approximately the age of 70 years in the multivariate analysis. Moreover, 70 years is a clinically relevant turning point.[14] The biologic reason for this breakpoint/transition phenomenon is not clear. Breast cancer has a somewhat different biologic behavior in older adult patients versus younger patients. Advancing age is associated with more favorable tumor biology.[2-4,6,20] On the other hand, there is a clear

 © 2009 by American Society of Clinical Oncology

Information downloaded from jco.ascopubs.org and provided by CATHARINA ZIEKENHUIS on June 3, 2009 from 137.56.139.73.
Copyright © 2009 by the American Society of Clinical Oncology. All rights reserved.

suppressed cellular immunity in older adults,[21] and the most important determinant is age as such rather than age-associated diseases.[22] The number of tumor-infiltrating lymphocytes in breast cancer decreases with age.[23] Moreover, patients with breast cancer with high cell-mediated immunity to tumor-associated antigens have a better prognosis than those with low immunity, indicating that immunologically unreactive patients are at risk for disease recurrence.[24] Thus altered or decreased immunologic function could result in decreased defense against invasion and increased nodal metastasizing, in a subset of breast tumors in older adults. In the large Surveillance, Epidemiology, and End Results database,[3] breast cancer survival in older women seemed to be similar to survival in the general population irrespective of disease status. It might well be that there is a balance in older adults between the less aggressive biologic phenotype on the one hand and the decreased immunologic defense on the other hand, resulting in status quo on the level of breast cancer survival. Immunologic and microarray studies might further help in elucidating age-related differences in tumor behavior and immunology. This might ultimately lead to better insight in tumor biology of breast cancer in general and reveal new opportunities for directing anticancer strategies.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Conception and design: Hans Wildiers, Ben Van Calster, Patrick Neven
Administrative support: Marie-Rose Christiaens
Provision of study materials or patients: Hans Wildiers, Lonneke V. van de Poll-Franse, Wouter Hendrickx, Ann Smeets, Robert Paridaens, Karen Deraedt, Karin Leunen, Caroline Weltens, Sabine Van Huffel, Marie-Rose Christiaens, Patrick Neven
Collection and assembly of data: Hans Wildiers, Ben Van Calster
Data analysis and interpretation: Hans Wildiers, Ben Van Calster, Jo Røislien
Manuscript writing: Hans Wildiers, Ben Van Calster, Jo Røislien
Final approval of manuscript: Hans Wildiers, Ben Van Calster, Lonneke V. van de Poll-Franse, Wouter Hendrickx, Jo Røislien, Ann Smeets, Robert Paridaens, Karen Deraedt, Karin Leunen, Caroline Weltens, Sabine Van Huffel, Marie-Rose Christiaens, Patrick Neven

REFERENCES

1. Gennari R, Curigliano G, Rotmensz N, et al: Breast carcinoma in elderly women: Features of disease presentation, choice of local and systemic treatments compared with younger postmenopausal patients. Cancer 101:1302-1310, 2004
2. Clark GM: The biology of breast cancer in older women. J Gerontol 47:19-23, 1992
3. Diab SG, Elledge RM, Clark GM: Tumor characteristics and clinical outcome of elderly women with breast cancer. J Natl Cancer Inst 92:550-556, 2000
4. Daidone MG, Coradini D, Martelli G, et al: Primary breast cancer in elderly women: Biological profile and relation with clinical outcome. Crit Rev Oncol Hematol 45:313-325, 2003
5. Bonnier P, Romain S, Charpin C, et al: Age as a prognostic factor in breast cancer: Relationship to pathologic and biologic features. Int J Cancer 62: 138-144, 1995
6. Fisher CJ, Egan MK, Smith P, et al: Histopathology of breast cancer in relation to age. Br J Cancer 75:593-596, 1997
7. Singh R, Hellman S, Heimann R: The natural history of breast carcinoma in the elderly: Implications for screening and treatment. Cancer 100:1807-1813, 2004
8. Ezzat A, Raja MA, Zwaan F, et al: The lack of age as a significant prognostic factor in non-metastatic breast cancer. Eur J Surg Oncol 24:23-27, 1998
9. Molino A, Giovannini M, Auriemma A, et al: Pathological, biological and clinical characteristics, and surgical management, of elderly women with breast cancer. Crit Rev Oncol Hematol 59:226-233, 2006
10. Neven P, Van Calster B, Van den Bempt I, et al: Age interacts with the expression of steroid and HER-2 receptors in operable invasive breast cancer. Breast Cancer Res Treat 110:153-159, 2008
11. Cleveland WS, Devlin SJ, Grosse E: Regression by local fitting: Methods, properties, and computational algorithms. J Econom 37:87-114, 1988
12. Schafer JL: Multiple imputation: A primer. Stat Methods Med Res 8:3-15, 1999
13. Little RJA, Rubin DB: Statistical Analysis with Missing Data (ed 2). New York, NY, John Wiley, 2002
14. Balducci L, Yates J: General guidelines for the management of older patients with cancer. Oncology (Huntingt) 14:221-227, 2000
15. Pastor-Barriuso R, Guallar E, Coresh J: Transition models for change-point estimation in logistic regression. Stat Med 22:1141-1162, 2003
16. Burnham KP, Anderson DR: Model Selection and Inference: A Practical Information-Theoretic Approach. New York, NY, Springer, 1998
17. Yancik R, Wesley MN, Ries LAG, et al: Effect of age and comorbidity in postmenopausal breast cancer patients aged 55 years and older. JAMA 285:885-892, 2001
18. Martelli G, Boracchi P, De Palo M, et al: A randomized trial comparing axillary dissection to no axillary dissection in older patients with T1N0 breast cancer: Results after 5 years of follow-up. Ann Surg 242:1-6, 2005
19. Holmberg SB, Crivelarri D, Zahrieh D, et al: A randomized trial comparing axillary clearance versus no axillary clearance in older patients (≥ 60 years) with breast cancer: First results of International Breast Cancer Study Group Trial 10-93. J Clin Oncol 22:4s, 2004 (suppl; abstr 505)
20. Pierga JY, Girre V, Laurence V, et al: Characteristics and outcome of 1755 operable breast cancers in women over 70 years of age. Breast 13: 369-375, 2004
21. Larbi A, Franceschi C, Mazzatti D, et al: Aging of the immune system as a prognostic factor for human longevity. Physiology (Bethesda) 23:64-74, 2008
22. Goodwin JS, Searles RP, Tung KS: Immunological responses of healthy elderly population. Clin Exp Immunol 48:403-410, 1982
23. Zavagno G, Meggiolaro F, Pluchinotta A, et al: Influence of age and menopausal status on pathologic and biologic features of breast cancer. Breast 9:320-328, 2000
24. Mccoy JL, Rucker R, Petros JA: Cell-mediated immunity to tumor-associated antigens is a better predictor of survival in early stage breast cancer than stage, grade or lymph node status. Breast Cancer Res Treat 60:227-234, 2000

Appendix

Details on the Methodology of the Multivariate Model on the Leuven Database

The multivariate analysis started with investigating the level of multicollinearity among the possible predictors. Next, for each predictor, we checked the assumption of linearity in the logit conditional on the other possible predictors (Hosmer DW, Lemeshow S. Applied logistic regression (ed 2). New York, NY, Wiley, 2000). Transformations, or other adjustments to predictors, were made if this assumption was violated. For variable selection, we focused on the Akaike and Bayesian Information Criteria (AIC, BIC)[16] and on the likelihood ratio *P* value and odds ratio (OR) for each variable. The AIC and BIC are model selection statistics that penalize for the number of predictors used. The AIC has the tendency

© 2009 by American Society of Clinical Oncology

Information downloaded from jco.ascopubs.org and provided by CATHARINA ZIEKENHUIS on June 3, 2009 from 137.56.139.73.
Copyright © 2009 by the American Society of Clinical Oncology. All rights reserved.

to select too many inputs, whereas the BIC has the tendency to select too few.[16] Therefore, instead of leaning on just one of these criteria, we monitored both, while simultaneously evaluating whether the resulting model seemed plausible from a clinical point of view.

Thereafter, the need to include interactions was investigated. However, caution on the inclusion of interactions is advised in statistical literature (Hosmer DW, Lemeshow S. Applied logistic regression (ed 2). New York, NY, Wiley, 2000; Harrell FE, Lee KL, Mark DB. Stat Med 15:361-387, 1996). Because of the large sample size and the extra risk of overfitting for complex terms (such as interactions), we were prudent concerning the inclusion of interaction terms. We investigated all two-way interactions one by one, as well as through stepwise and manual selection procedures. The main criterion for evaluating interactions was the BIC, as this is a strict statistic for model selection. A second criterion was the likelihood ratio *P* value for each interaction.

Finally, influence diagnostics were investigated to check for outliers and the extent to which they influence the model's results. All statistical tests were two-sided and were performed with SAS version 9.1 (SAS Institute Inc, Cary, NC).

Details on the Multiple Imputation Procedure for Missing Values in the Eindhoven Cancer Registry Data Set

We imputed the missing values in the Eindhoven Cancer Registry (ECR) data set five times. Each imputation was based on an iterative algorithm along the lines of, but not identical to, Siddique and Belin (Stat Med 27:83-102, 2008). The variables considered were both core variables and auxiliary variables. The core variables were the number of affected lymph nodes (0, one to three, four or more), age, size, estrogen receptor, progesterone receptor, HER-2, and grade. The auxiliary variables were histology (ductal, lobular, mixed, other; no missing values), TNM staging for tumor (1 to 4; 11.8% missing values), and inclusion year (2005 or 2006; no missing values). Auxiliary variables are used to improve the quality of the imputations (Rubin DB. J Am Stat Assoc 91:473-489, 1996; Schafer JL. Analysis of incomplete multivariate data. London, United Kingdom, Chapman and Hall, 1997). A completed data set was constructed as follows. In the first step, initial imputations for missing values were obtained using the expectation-maximization algorithm, with the assumption that the joint distribution of all variables involved is multivariate normal. In the second step, all variables with missing values were imputed in a predefined random order (this is the first iteration). Each variable was imputed using the information from all remaining variables. If one of the remaining variables had missing values too, the most recent imputations of these missing values were used such that all patients could be used. The continuous variable size was imputed using predictive mean matching (PMM) regression (Schenker N, Taylor JMG. Comput Stat Data Anal 22:425-446, 1996); categoric or ordinal variables were imputed using logistic regression. (A short description of how these methods work is given at the end of this appendix.) In the third step, the second step was repeated a number of times (new iterations). We used 80 iterations to decrease the effect of the initial imputations and variable imputation order on the final imputations. In addition, we observed that 80 iterations provided fairly stable results. We repeated the whole procedure five times to produce five completed data sets. Typically, three to 10 completed data sets suffice for MI.[12]

The imputation model should be rich (hence the use of auxiliary variables) and should include any association that is the subject of investigation in subsequent analyses (Schafer JL. Analysis of incomplete multivariate data. London, United Kingdom, Chapman and Hall, 1997). In the Leuven data, we detected a piecewise effect of age on lymph node involvement, with the effect of age changing at approximately 70 years of age (see Results). In addition, for women older than 70 years, the effect of age interacted with tumor size. Because we aimed to validate these effects using the ECR data, we had to include these associations in the imputation model. Therefore, we performed the MI procedure separately for women up to 70 years and women older than 70 years. Furthermore, for the latter group of patients, the PMM regression model to impute size included the interaction effect of age and number of affected lymph nodes, and the logistic regression model to impute the number of affected lymph nodes included the interaction effect of age and size.

The PMM regression method to impute size starts with fitting an ordinary least squares regression model. Then, new values for the model parameters are drawn from their posterior distribution. Using these new values, a predicted value for size is constructed for each patient. For a patient *x* with missing tumor size, we select the five patients with available tumor size whose predicted value for size is closest to that for *x*. One of the five closest matches is selected at random, and its tumor size is used to impute the missing size of patient *x*.

Similarly, the logistic regression method to impute categoric/ordinal variables starts with fitting a logistic regression model. Then, new values for the model parameters are drawn from their posterior distribution. The new values are used to compute the probability of each value of the target variable. One specific value is randomly imputed using these probabilities. Both methods were implemented using PROC MI in SAS.

Combining the results of the five completed data sets was done using PROC MIANALYZE in SAS.

Multivariate Logistic Regression Model on the Leuven Database

The probability of lymph node involvement was estimated as $1/(1 + e^{-z})$, where

$$z = -1.6616 - 0.01433 \cdot \text{age} + 0.1217 \cdot (\text{age} - 70) \cdot \text{age group} + 0.04677 \cdot \text{size} - 0.002516 \cdot (\text{age} - 70) \cdot \text{age group} \cdot \text{size} + 0.2443 \cdot \text{grade} + 0.7386 \cdot \text{triple positivity (TP)} \quad (1)$$

Age group equals 0 if the patient is up to 70 years and 1 if the patient is older than 70 years. This allowed the reformulation of the model for younger and older patients separately, by replacing age group with 0 or 1, respectively:

$$z_{\text{young}} = -1.6616 - 0.01433 \cdot \text{age} + 0.04677 \cdot \text{size} + 0.2443 \cdot \text{grade} + 0.7386 \cdot \text{TP} \quad (2)$$

$$z_{\text{old}} = -10.1806 + 0.1074 \cdot \text{age} + 0.2229 \cdot \text{size} - 0.002516 \cdot \text{age} \cdot \text{size} + 0.2443 \cdot \text{grade} + 0.7386 \cdot \text{TP} \quad (3)$$

Multivariate Logistic Regression Model on the ECR

The probability of lymph node involvement was estimated as $1/(1+e^{-z})$, where:

Information downloaded from jco.ascopubs.org and provided by CATHARINA ZIEKENHUIS on June 3, 2009 from 137.56.139.73.
Copyright © 2009 by the American Society of Clinical Oncology. All rights reserved.

$$z = -0.8877 - 0.01947 \cdot \text{age} + 0.08058 \cdot (\text{age} - 70) \cdot \text{age group} + 0.06342 \cdot \text{size} - 0.002050 \cdot (\text{age} - 70) \cdot \text{age group} \cdot \text{size} + 0.1706 \cdot \text{grade} + 0.05518 \cdot \text{TP} \quad (4)$$

Reformulating the model for younger and older patients separately gave:

$$z_{\text{young}} = -0.8877 - 0.01947 \cdot \text{age} + 0.06342 \cdot \text{size} + 0.1706 \cdot \text{grade} + 0.05518 \cdot \text{TP} \quad (5)$$

$$z_{\text{old}} = -6.5283 + 0.06111 \cdot \text{age} + 0.2069 \cdot \text{size} - 0.002050 \cdot \text{age} \cdot \text{size} + 0.1706 \cdot \text{grade} + 0.05518 \cdot \text{TP} \quad (6)$$

© 2009 by American Society of Clinical Oncology

Information downloaded from jco.ascopubs.org and provided by CATHARINA ZIEKENHUIS on June 3, 2009 from 137.56.139.73.
Copyright © 2009 by the American Society of Clinical Oncology. All rights reserved.