VOLUME 27 · NUMBER 7 · MARCH 1 2009

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Lymph Node Ratio as an Alternative to pN Staging in Node-Positive Breast Cancer

Vincent Vinh-Hung, Helena M. Verkooijen, Gerald Fioretta, Isabelle Neyroud-Caspar, Elisabetta Rapiti, Georges Vlastos, Carole Deglise, Massimo Usel, Jean-Michel Lutz, and Christine Bouchardy

From the Geneva Cancer Registry, Institute for Social and Preventive Medicine, Geneva University; Senology and Surgical Gynecologic Oncology Unit, Department of Gynecology and Obstetrics, Hôpitaux Universitaires de Genève, Geneva, Switzerland; Universitair Ziekenhuis Brussel, Jette, Belgium; and Department of Community Occupational and Family Medicine, National University of Singapore, Singapore.

Submitted June 17, 2008; accepted September 30, 2008; published online ahead of print at www.jco.org on January 21, 2009.

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

Corresponding author: Vincent Vinh-Hung, MD, PhD, Oncology Center, Universitair Ziekenhuis Brussel, 101 Laarbeeklaan, 1090 Jette, Belgium; e-mail: conrvhgv@uzbrussel.be.

0732-183X/09/2707-1062/$20.00

DOI: 10.1200/JCO.2008.18.6965

ABSTRACT

Purpose
In the current pTNM classification system, nodal status of breast cancer is based on the number of involved lymph nodes and does not account for the total number of lymph nodes removed. In this study, we assessed the prognostic value of the lymph node ratio (LNR; ie, ratio of positive over excised lymph nodes) as compared with pN staging and determined its optimal cutoff points.

Patients and Methods
From the Geneva Cancer Registry, we identified all women diagnosed with node-positive breast cancer between 1980 and 2004 (n = 1,829). The prognostic value of LNRs was calculated for values ranging from 0.05 to 0.95 by Cox regression analysis and validated by bootstrapping. Based on maximum likelihood, we identified cutoff points classifying women into low-, intermediate-, and high-risk LNR groups.

Results
Optimal cutoff points classified patients into low- (≤ 0.20), intermediate- (> 0.20 and ≤ 0.65), and high-risk (> 0.65) LNR groups, corresponding to 10-year disease-specific survival rates of 75%, 63%, and 40%, and adjusted mortality risks of 1 (reference), 1.78 (95% CI, 1.46 to 2.18), and 3.21 (95% CI, 2.54 to 4.06), respectively. In contrast to LNR risk categories, survival curves of pN2 and pN3 crossed after 15 years, and their adjusted mortality risks showed overlapping CIs: 2.07 (95% CI, 1.69 to 2.53) and 2.84 (95% CI, 2.23 to 3.61), respectively.

Conclusion
LNR predicts survival after breast cancer more accurately than pN classification and should be considered as an alternative to pN staging.

J Clin Oncol 27:1062-1068. © 2009 by American Society of Clinical Oncology

INTRODUCTION

Axillary lymph node involvement and the number of involved axillary lymph nodes are among the most important prognostic factors in breast cancer.[1,2] The number of involved axillary nodes has been incorporated into routine clinical decision making,[3-5] and according to the sixth edition of the American Joint Committee on Cancer/International Union Against Cancer staging system, patients with one to three positive axillary lymph nodes are classified as having pN1 disease, patients with four to nine positive axillary lymph nodes are classified as having pN2 disease, and patients with 10 or more positive axillary lymph nodes are classified as having pN3 disease.[6-8]

The number of involved lymph nodes identified depends on the number of lymph nodes removed and examined, which in itself depends on the surgical and pathologic procedure. In these cases with few nodes removed, patients cannot be classified as having pN3 disease, which can affect comparisons between institutions where the practices of axillary dissection differ. To improve the prognostication system, one would intuitively take not only the number of positive lymph nodes, but also the number of nodes examined into account.[9] The lymph node ratio (LNR), defined as the number of involved nodes divided by the number of lymph nodes examined, standardizes against the variability of nodal assessment and was found to improve prognostic information when compared with the number of involved nodes. Woodward et al[10] conducted a systematic review of 24 articles published between 1994 and 2005 totaling 32,299 patients, of whom 3,565 were from four randomized trials, and 18,038 patients were from a Surveillance, Epidemiology, and End Results study. The LNR was confirmed to be superior to the number of involved nodes as a prognostic indicator. A subsequent study

comparing patients included in prospective trials showed that the LNR improved the comparability of institutions.[11] However, to date, there has been no formal proposal toward using the LNR as an alternative to the current pN staging. Moreover, a robust and reproducible categorization can be important to identify subgroups of patients who might better respond to a treatment or to efficiently plan prospective trials. Hence we will address in this article whether patients with breast cancer can be classified into meaningful risk categories based on LNR, by comparing it with pN staging.

PATIENTS AND METHODS

We used data from the Geneva cancer registry, which records information on all incident cancer cases that occur in the canton (approximately 420,000 inhabitants). The registration is based on several sources of information and is accurate, as attested by its low percentage (< 2%) of cases recorded from death certificates only.[12]

Information recorded for each patient includes sociodemographic data, diagnostic circumstances, diagnostic modalities, histologic features of the tumor, treatment during the first 6 months after diagnosis, survival, and cause of death. In addition to passive follow-up (routine examination of death certificates and hospital records), the registry regularly assesses survival through an active follow-up performed routinely each year at the Cantonal Population Office, which is in charge of the registration of the resident population. For all deceased patients, the registry's medical staff systematically consults medical files, writes to practitioners to assess the exact cause of death, and codes the cause according to the WHO classification.[13]

This study included all female residents of the Swiss canton of Geneva with a nonmetastatic primary invasive breast carcinoma diagnosed between 1980 and 2004 (n = 6,936). We selected women who underwent axillary lymph node dissection, in whom the total number of nodes examined was mentioned in the pathology report (n = 5,053), and who presented with one or more involved (ie, positive) lymph nodes (n = 1,924 after exclusion of 2,954 node-negative cases and 175 cases with an unknown number of positive nodes). We excluded records in which tumor size was not reported (n = 95). The final number available was 1,829 patients, representing the study population. A separate coding for the sentinel node procedure was introduced in 2001; the selection did not retain 118 cases of positive sentinel nodes because of missing total number of nodes.

We considered sociodemographic characteristics, tumor characteristics, and treatment covariates for the models. Sociodemographic variables were age (continuous), and year of diagnosis (continuous). Economic status was based on the woman's most recent occupation, or that of the spouse for the unemployed, and was categorized as high (professionals, executives, and administrators) versus lower classes. Histologic features considered were grade (poorly or undifferentiated *v* others), log of tumor size in centimeters (continuous), number of nodes removed (continuous), number of positive nodes (continuous), and LNR (continuous). Adjuvant treatments considered were radiotherapy of breast or chest wall (yes/no), chemotherapy (yes/no), and hormone therapy (yes/no). End point for the survival analyses was death from breast cancer. Patients who died from other causes were censored at the time of death. Patients who left the canton were censored at the last date of known residence.

The analysis was done in two stages. In the first stage, we evaluated the prognostic value of LNR, adjusting for other above-mentioned covariates significantly associated with breast cancer mortality. With Cox proportional hazards analyses, we calculated breast cancer–specific mortality risks. We applied the Akaike Information Criteria (AIC) to identify the covariates that were the most significantly associated with mortality and to prevent overadjustment.[14] We used no other criteria. In the second stage, after ascertaining that the LNR was indeed significantly associated with breast cancer mortality, we proceeded to determine the most appropriate cut point for categorizing LNR as high risk, medium risk, and low risk. For this, we recomputed the likelihood associated with all possible pairs of LNR cutoffs (dividing patients into high-, medium-, and low-risk categories) ranging from 0.05 to 0.95 at intervals of 0.05.[15] We recorded the differences between the likelihoods of the cutoff models (where the LNR was categorized by the cut points) and the AIC model (where the LNR was modeled as a continuous covariate). We retained the pair of cutoffs associated with the least negative difference in likelihoods (ie, the pair causing the least loss of information resulting from categorization).

To estimate the stability of the results, we used a bootstrap procedure,[14,16] which applies the proportional hazards computations to full random samples with replacement of the patients. We ran 10,000 iterations of the procedure.

We also examined the impact of specifying a minimum number of lymph nodes excised on the accuracy of the LNR by recomputing the hazard ratios for different minimum numbers of lymph nodes excised.

The study was not submitted to an internal review board but was developed through consensus meetings within the registry. All statistical computations used R version 2.6.1.[17] The AIC procedure used the MASS library.[18] The likelihood profiles were smoothed using the "mgcv" package.[19]

Table 1. Characteristics of Patients With Lymph Node–Positive Breast Cancer

Characteristic	No. of Patients (N = 1,829)	%
Age, years		
Median	58.9	
Lower-upper quartiles	49.8-68.9	
< 50	465	25.4
≥ 50	1,364	74.6
Year of diagnosis		
1980-1984	301	16.5
1985-1989	323	17.7
1990-1994	381	20.8
1995-1999	473	25.9
2000-2004	351	19.2
High socioeconomic class	185	10.1
Medial tumor location	240	13.1
High histologic grade	404	22.1
Tumor size, cm		
Median	2.5	
Lower-upper quartiles	2-3	
≤ 2	806	44.1
> 2 cm	1,023	55.9
No. of lymph nodes removed		
Median	14	
Lower-upper quartiles	10-19	
1-3	30	1.6
4-6	73	4.0
6-9	248	13.6
≥ 10	1478	80.8
No. of positive lymph nodes		
Median	2	
Lower-upper quartiles	1-5	
1-3, pN1	1,178	64.4
4-9, pN2	429	23.5
≥ 10, pN3	222	12.1
Lymph node ratio		
Median	0.18	
Lower-upper quartiles	0.09-0.40	
≤ 0.20	1,024	56.0
> 0.20 and ≤ 0.65	553	30.2
> 0.65	252	13.8
Adjuvant treatment		
No adjuvant treatment	190	10.4
Radiotherapy	1,238	67.7
Chemotherapy	1,051	57.5
Endocrine therapy	843	46.1

Table 2. Prognostic Factors* of Breast Cancer Mortality Among Patients With Lymph Node–Positive Breast Cancer

Variable	Hazard Ratio	95% CI	*P*
Age at diagnosis, continuous	1.02	1.01 to 1.02	.0001
Year of diagnosis, continuous	0.95	0.94 to 0.97	< .0001
High socioeconomic class *v* other	0.74	0.53 to 1.04	.0819
Tumor medial location *v* other	1.22	0.96 to 1.56	.1031
High histologic grade *v* other	1.64	1.33 to 2.02	< .0001
Log tumor size (cm), continuous	1.86	1.56 to 2.23	< .0001
Radiotherapy *v* no	0.70	0.58 to 0.84	.0001
Chemotherapy *v* no	0.79	0.64 to 0.96	.0198
Endocrine therapy *v* no	0.75	0.60 to 0.94	.0105
Lymph node ratio, continuous	4.51†	3.37 to 6.04	< .0001

*Cox proportional hazards model selected by Akaike Information Criteria stepwise regression; only deaths from breast cancer are considered. Hazard ratios are adjusted for all other factors listed in the table.
†Caveat about the interpretation (see text discussion).

RESULTS

The median age of the 1,829 patients was 58.9 years (Table 1). The number of patients in the last 5 years decreased as a result of the separate coding of sentinel node procedure and corresponding missing cases. A large majority of patients (81%) had at least 10 axillary lymph nodes removed. The median number of involved nodes was two (range, one to 32 nodes), and the mean LNR was 0.18 (range, 0.016 to 1.000). Most patients (90%) had received radiotherapy, chemotherapy, or hormone therapy, alone or in combination.

In the first stage of multivariate analyses, patients' age, year of diagnosis, grade, size, radiation treatment, chemotherapy, endocrine therapy, and LNR were independent and significant prognostic factors of breast cancer mortality (Table 2). LNR as a continuous covariate was the most important prognostic factor, with a 4.51-fold increased relative risk, with LNR expressed on a fractional scale from 0 to 1, corresponding to a relative breast cancer mortality increase of 1.5% per 1% involved nodes, with LNR expressed on a percentage scale. Socioeconomic class and tumor location in inner quadrants were not statistically significantly associated with breast cancer mortality (Table 2). Nevertheless, these two variables contributed to improve the AIC and were therefore retained in the model. The number of positive nodes was not retained. Examination of 10,000 resamplings of the data with reiterations of the AIC selection showed that the LNR was retained in 99.98% of the random samples, whereas the number of removed nodes and the number of positive nodes were retained respectively in only 76.38% and 26.39% of the random samples.

In the second stage, we used the multivariate model as identified in Table 2, but iteratively replacing the continuous LNR with different pairs of categorized LNRs. Figure 1 summarizes the distribution histograms of the LNR cutoff points that were associated with the least negative change of maximum partial likelihood based on 10,000 bootstrap iterations. The figure indicates the number of times any given cutoff point was retained. The distribution histograms show a sharply defined lower LNR cutoff point of 0.20 (mean, 0.21; standard deviation, 0.04; 95% CI, 0.15 to 0.30; Fig 1A) and an upper LNR cutoff point of 0.65 (mean, 0.67; standard deviation, 0.13; 95% CI, 0.30 to 0.95; Fig 1B). The presence of a separate smaller peak (Fig 1B) indicates that in some random samples, the upper cutoff point was very close to the lower cutoff point, but this occurred only in 5% of the samples.

Using the pair of cutoff points of 0.20 and 0.65, we classified patients with LNR $\leq$ 0.20 into low-risk, LNR more than 0.20 and $\leq$ 0.65 into intermediate-risk, and LNR more than 0.65 into high-risk categories. Figure 2 shows the univariate Kaplan-Meier survival estimates according to risk groups defined by pN staging (Fig 2A) or defined by the LNR (Fig 2B). The breast cancer–specific survival at 10 years for patients with low-risk, intermediate-risk, and high-risk LNR were 75.2%, 63.3%, and 39.6%, respectively (log-rank $\chi^2 = 155.7$). In addition, the pN classification showed an imbalance in prognostic separation, with the pN2 and pN3 survival curves crossing after 15 years, whereas the nodal-ratio curves remained separated, even with a follow-up exceeding 20 years (Fig 2). In multivariate analysis, compared with patients within the low LNR risk group, the adjusted hazard ratio of breast cancer mortality risk was 1.78 (95% CI, 1.46 to 2.18) for patients in the intermediate LNR risk group and 3.21 (95% CI, 2.54 to 4.06) for patients within the high LNR risk group (Table 3). By comparison, the survival rates for patients with pN1, pN2, and pN3 disease were 75.8%, 54.4%, and 42.7%, respectively ($\chi^2 = 150.5$). Compared with patients with pN1 stage, the adjusted hazard ratio of breast cancer mortality was 2.07 (95% CI, 1.69 to 2.53) for patients with pN2 stage and 2.84 (95% CI, 2.23 to 3.61) for patients with pN3

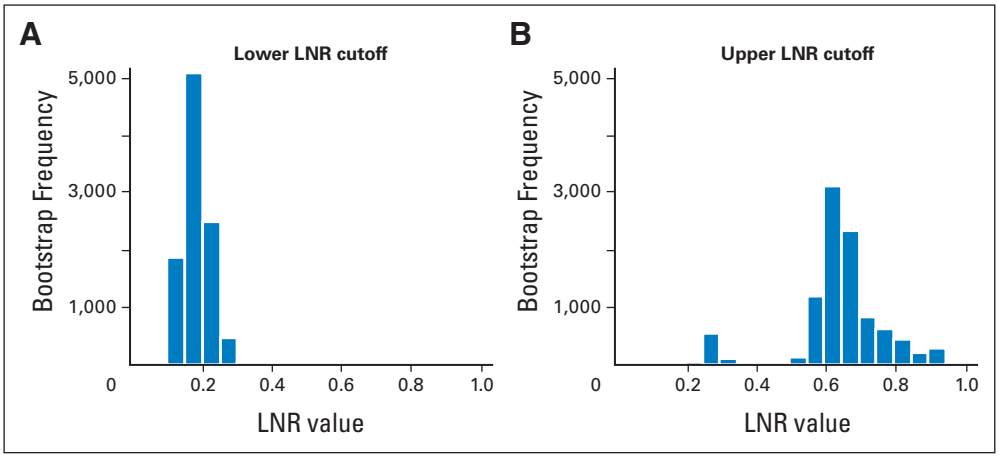

Fig 1. Distribution of bootstrap lymph node ratio (LNR) cutoff points: (A) lower cutoff point; (B) upper cutoff point.

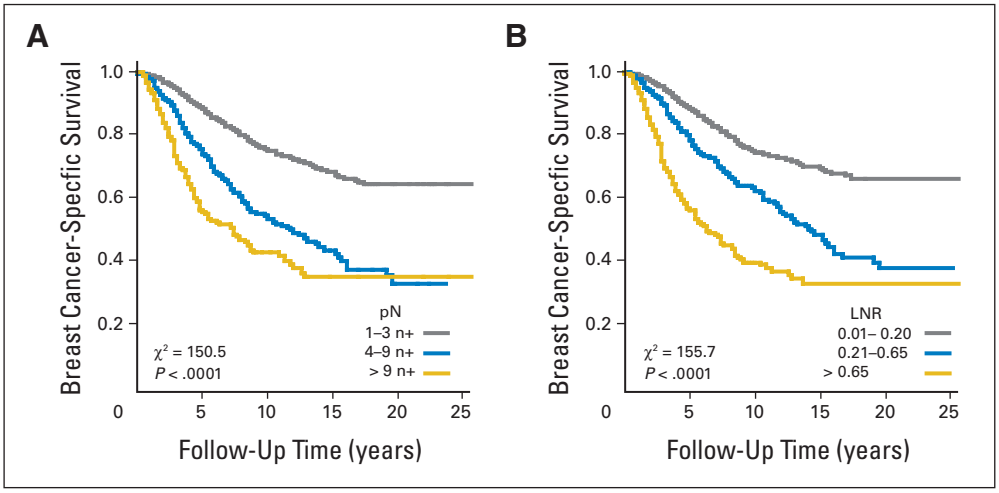

Fig 2. Kaplan-Meier survival estimates according to risk groups: (A) risk groups defined by pN; (B) risk groups defined by lymph node ratio (LNR).

stage (Table 3). The CIs of the pN2 and pN3 hazard ratios overlapped, in contrast with the LNR hazard ratios, which did not overlap.

We looked at the impact of specifying a minimum number of excised lymph nodes on the reliability on the pN and the LNR hazard ratios, respectively. Figure 3 shows the hazard ratios from the same multivariate model as in Table 3, recomputed each time by successively excluding patients with fewer than one—then two, three, and so on—lymph nodes excised. The overlapping of the pN CIs was not improved by specifying a minimum number of lymph nodes to be removed (Fig 3A). The LNR CIs remained separated, up to 13 excised lymph nodes (Fig 3B). In both pN and LNR cases, the width of the CIs increased, indicating loss of precision with more stringent minimum specification.

DISCUSSION

In this study, the LNR was one of the most important prognostic factors of breast cancer mortality. The LNR provided a better classification of patients' prognostic risk profile than the pN classification system, in particular after 10 years. Also, this study proposes, for the first time, a categorization of the LNR validated by bootstrap resampling among a population-based cohort of women with lymph node–positive breast cancer.

Table 3. Effect of Lymph Node Ratio and pN Classification on Breast Cancer Mortality Among Patients With Lymph Node–Positive Breast Cancer

Variable	Hazard Ratio*	95% CI	*P*
Lymph node ratio			< .0001
Low, ≤ 0.20	1	Reference	
Medium, > 0.20 and ≤ 0.65	1.78	1.46 to 2.18	
High, > 0.65	3.21	2.54 to 4.06	
pN			< .0001
pN1	1	Reference	
pN2	2.07	1.69 to 2.53	
pN3	2.84	2.23 to 3.61	

*Cox proportional hazards model; only deaths from breast cancer are considered. Hazard ratios are adjusted for age, year of diagnosis, socioeconomic class, tumor location, histologic grade, tumor size, radiotherapy, chemotherapy, and endocrine therapy.

There is growing evidence establishing the prognostic role of the LNR in breast cancer.[9-11,20-22] The importance of the LNR has been shown for many cancer sites from the Surveillance, Epidemiology, and End Results population data: esophagus,[23] vulva,[24] colon,[25] corpus uteri,[26] or from case series, gastric,[27,28] pancreas,[29] and bladder,[30] indicating the potential of the LNR as the basis of a staging system. However, there is no clear consensus about the cutoff points that would be required for a staging classification. As reviewed by Woodward et al,[10] the cutoff points used by different authors to classify patients with breast cancer into more than two risk groups differ widely: 0.10/0.50,[31] 0.50/0.75,[32] 0.33/0.67,[33,34] 0.17/0.43/0.85,[35] 0.25/0.50,[36] 0.10/0.50,[20] and 0.25/0.50/0.80.[37] We tried to identify cutoff points that should be robust against the variability of the data. Hence we chose the bootstrap resampling method, which does not necessarily rely on *P* values or distributional assumptions.[38,39] We found that the lower cutoff point was sharply defined with a narrow CI, suggesting high reliability. The value of 0.20 closely matches the most recent reports from other centers.[10,11,21,22] Interestingly, the upper cutoff point was less sharply defined. The wider CI is attributable to the smaller number of patients with a high LNR. But it also suggests an increased heterogeneity among patients with more extensive nodal involvement.[40]

Categorization of a continuous variable (in our case, the LNR) has been shown to be associated with several problems, including loss of information,[38,41,42] inflation of type I error rate,[43] increase in variance of estimated hazard ratios, loss of power, and decrease in efficiency of survival analysis[44,45] and can even mislead one into concluding that a second unimportant variable is important.[46] In the present study, we built the model with LNR as a continuous variable. Categorization was done afterward. In examining the output of the bootstrap iterations, which generated 1.67 million comparisons between the continuous LNR and the various categorized LNR models, we found that the categorized models were poorer than the continuous models in 89.8% of the comparisons, confirming that the LNR should be maintained as a continuous covariate for modeling purposes.

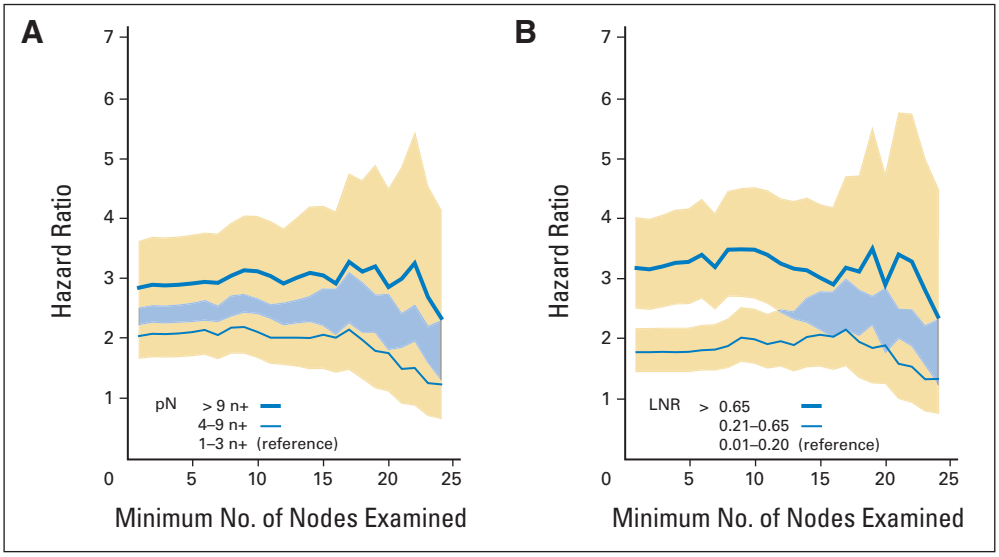

Fig 3. Multivariate breast cancer mortality hazard ratios of the (A) pN risk groups and (B) lymph node ratio (LNR) risk groups as a function of a minimum number of nodes examined. Gray bands are 95% CIs; darker bands show their overlap.

However, these reasons do not preclude the need for categorization, which can arise because of the complexity of underlying computations, as discussed hereafter. We used a scale from 0 to 1 for the LNR, which gave a hazard ratio of 4.51. The LNR is linear in the logarithm of the hazard ratio. To compute, for example, the relative risk of breast cancer death associated with the involvement of one node of five nodes total, as compared with no nodal involvement, we need to convert the hazard ratio into its corresponding regression coefficient, 1.506 = Log(4.51), multiply this coefficient by 0.20 (= 1/5), which gives 0.301 (= 1.506 × 0.20), then exponentiate to obtain a hazard ratio of 1.35 = Exp(0.301). This need to switch back and forth between logarithm and exponentiation is awkward for daily clinical practice, it cannot be obviated by changing the scale, and it can be prone to misinterpretation. By contrast, hazard ratios of categorized LNR such as those shown in Table 3 are immediately readable.

Our categorization of the LNR showed a clear advantage over pN staging: the 3.21 hazard ratio of the high-risk LNR indicates a separation between the high- and low-risk group (3.21 − 1 = 2.21) that is wider than the separation between pN3 (hazard ratio = 2.84) and pN1 (2.84 − 1 = 1.74). Moreover, the intermediate-risk LNR was truly intermediate; its CI overlapped neither the low risk nor the high-risk LNRs, whereas the pN2 and pN3 CIs overlapped (Table 3). Thus in multivariate analyses, classification using the LNR provided well-balanced nonoverlapping risk groups, whereas classification using pN provided poorly separated risk groups with overlapping hazard ratios.

The advantage of the LNR classification over pN staging was also apparent in unadjusted survival analysis. The log-rank χ^2 associated with the LNR was larger than that of pN (Fig 2), indicating a higher statistical significance. The LNR provided a balanced separation between the survival curves. The survival curves of the three LNR curves did not cross, even with a follow-up exceeding 15 years. By contrast, the survival curves of the pN risk groups were unbalanced. The pN2 and pN3 curves were close to one another and crossed, indicating graphically a poorer separation between intermediate- and high-risk groups (Fig 2A).

The question arises whether LNR-based classification should replace pN classification. For homogeneous data—if, for example, all patients underwent the same extensive axillary dissection—the distinction between a number-based and a ratio-based staging would disappear, and there would be no advantage of replacing the pN with an LNR-based classification. However, heterogeneity of lymph node examination is commonly encountered in daily practice; the LNR can be useful to address that heterogeneity. Our present population study found that the LNR improved over the pN. The improvement was observed in both multivariate and univariate analyses. In prospective randomized clinical trials, we believe that it would be more meaningful to balance patients allocation using the LNR-based risk groups rather than the pN. Taking into account that other authors have reported that the LNR also improved the comparison between centers,[11] our results argue that the LNR should be considered as an alternative to pN staging.

We acknowledge several limitations. We need additional studies on different populations and other health systems to validate our results and generalize our conclusions. The LNR assumes all lymph nodes are similarly examined, an unlikely condition. Our modeling did not account for the size of nodal metastases or factors like extracapsular or vascular invasion. We did not address supraclavicular or internal mammary chain involvement. The analyses were restricted to patients with measurable primary tumors. Furthermore, as sentinel node biopsy is progressively replacing axillary lymph node dissection, the value of the LNR becomes questionable.

Strengths of this study are the population basis and accurate follow-up data. Regarding the sentinel node biopsy, a growing literature indicates that the ratio of involved sentinel nodes and the proportion of sentinel nodes replaced by metastasis are important predictors of nonsentinel node involvement.[47-56] Our results blend with these concepts of tumor load in the sentinel nodes. We believe that LNR-based staging might provide a smooth transition toward biomarker-based staging in the near future, such as hinted by the gene recurrence score[57] or circulating tumor cells.[58,59]

In conclusion, our study identifies two LNR cutoff points, 0.20 and 0.65, which define breast cancer prognosis more adequately than the pN categories. We argue that nodal ratios should be considered as an alternative to pN staging.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

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

AUTHOR CONTRIBUTIONS

Conception and design: Vincent Vinh-Hung
Financial support: Christine Bouchardy
Administrative support: Christine Bouchardy
Provision of study materials or patients: Gerald Fioretta, Isabelle Neyroud-Caspar, Christine Bouchardy
Collection and assembly of data: Gerald Fioretta, Isabelle Neyroud-Caspar, Carole Deglise, Christine Bouchardy
Data analysis and interpretation: Vincent Vinh-Hung, Helena M. Verkooijen, Gerald Fioretta, Elisabetta Rapiti, Georges Vlastos, Massimo Usel, Christine Bouchardy
Manuscript writing: Vincent Vinh-Hung, Helena M. Verkooijen, Isabelle Neyroud-Caspar, Elisabetta Rapiti, Georges Vlastos, Carole Deglise, Massimo Usel, Jean-Michel Lutz, Christine Bouchardy
Final approval of manuscript: Vincent Vinh-Hung, Helena M. Verkooijen, Gerald Fioretta, Isabelle Neyroud-Caspar, Elisabetta Rapiti, Georges Vlastos, Carole Deglise, Massimo Usel, Jean-Michel Lutz, Christine Bouchardy

REFERENCES

1. Fitzgibbons PL, Page DL, Weaver D, et al: Prognostic factors in breast cancer: College of American Pathologists Consensus Statement 1999. Arch Pathol Lab Med 124:966-978, 2000

2. Vinh-Hung V, Burzykowski T, Cserni G, et al: Functional form of the effect of the numbers of axillary nodes on survival in early breast cancer. Int J Oncol 22:697-704, 2003

3. Hudis CA, Seidman AD, Baselga J, et al: Sequential adjuvant therapy with doxorubicin/paclitaxel/cyclophosphamide for resectable breast cancer involving four or more axillary nodes. Semin Oncol 22:18-23, 1995

4. Gianni AM, Siena S, Bregni M, et al: Efficacy, toxicity, and applicability of high-dose sequential chemotherapy as adjuvant treatment in operable breast cancer with 10 or more involved axillary nodes: Five-year results. J Clin Oncol 15:2312-2321, 1997

5. Fountzilas G, Nicolaides C, Aravantinos G, et al: Dose-dense adjuvant chemotherapy with epirubicin monotherapy in patients with operable breast cancer and >/=10 positive axillary lymph nodes: A feasibility study. Oncology 55:508-512, 1998

6. Singletary SE, Allred C, Ashley P, et al: Revision of the American Joint Committee on Cancer staging system for breast cancer. J Clin Oncol 20:3628-3636, 2002

7. Greene FL, Page DL, Fleming ID, et al: AJCC Cancer Staging Handbook (6th ed). New York, NY, Springer Verlag, 2002, pp 255-281

8. Sobin LH, Wittekind Ch (eds): TNM Classification of Malignant Tumours (ed 6). NewYork, NY, Wiley, 2002, pp 131-141

9. Vinh-Hung V, Verschraegen C, Promish DI, et al: Ratios of involved nodes in early breast cancer. Breast Cancer Res 6:R680–R688, 2004

10. Woodward WA, Vinh-Hung V, Ueno NT, et al: Prognostic value of nodal ratios in node-positive breast cancer. J Clin Oncol 24:2910-2916, 2006

11. Truong PT, Woodward WA, Thames HD, et al: The ratio of positive to excised nodes identifies high-risk subsets and reduces inter-institutional differences in locoregional recurrence risk estimates in breast cancer patients with 1-3 positive nodes: An analysis of prospective data from British Columbia and the M. D. Anderson Cancer Center. Int J Radiat Oncol Biol Phys 68:59-65, 2007

12. Bouchardy C: Switzerland, Geneva, in Parkin DM, Whelan SL, Ferlay J, Raymond L, Young J (eds): Cancer Incidence in Five Continents (vol VII). Lyon, France, International Agency for Research on Cancer, 1997, pp 666-669

13. World Health Organization: ICD-O International Classification of Diseases for Oncology (ed 1). Geneva, Switzerland, World Health Organization, 1976

14. Tableman M, Kim JS: Survival Analysis Using S: Analysis of Time-to-Event Data. Boca Raton, FL, Chapman & Hall, 2003

15. Vinh-Hung V, Vlastos G, Fioretta G, et al: Nodal ratios as an alternative to pN-staging in node-positive breast cancer: A validation study. Breast Cancer Res Treat 106:S215, 2007 (suppl, abstr 5027)

16. Gong G: Cross-validation, the jackknife, and the bootstrap: Excess error estimation in forward logistic regression. J Am Stat Assoc 81:108-113, 1986

17. The R Foundation for Statistical Computing: The R project for statistical computing: The Comprehensive R Archive Network (CRAN) 2008. http://www.r-project.org/

18. Venables WN, Ripley BD: Modern Applied Statistics with S (ed 4). New York, NY, Springer-Verlag, 2002, pp 172-176

19. Wood SN: Generalized Additive Models: An Introduction with R. Boca Raton, FL, Chapman & Hall/CRC Press, 2006

20. Voordeckers M, Vinh-Hung V, Van de Steene J, et al: The lymph node ratio as prognostic factor in node-positive breast cancer. Radiother Oncol 70: 225-230, 2004

21. Kuru B: Prognostic significance of total number of nodes removed, negative nodes removed, and ratio of positive nodes to removed nodes in node positive breast carcinoma. Eur J Surg Oncol 32:1082-1088, 2006

22. Yildirim E, Berberoglu U: Lymph node ratio is more valuable than level III involvement for prediction of outcome in node-positive breast carcinoma patients. World J Surg 31:276-289, 2007

23. Greenstein AJ, Litle VR, Swanson SJ, et al: Prognostic significance of the number of lymph node metastases in esophageal cancer. J Am Coll Surg 206:239-246, 2008

24. Vlastos AT, Vlastos G, De Ridder M, et al: Prognostic importance of nodal involvement in vulvar carcinoma (VC). Int J Gynecol Cancer 16:732-733, 2006 (suppl 3, abstr 0469)

25. De Ridder M, Vinh-Hung V, Van Nieuwenhove Y, et al: Prognostic value of the lymph node ratio in node positive colon cancer. Gut 55:1681, 2006

26. Chan JK, Kapp DS, Cheung MK, et al: The impact of the absolute number and ratio of positive lymph nodes on survival of endometrioid uterine cancer patients. Br J Cancer 97:605-611, 2007

27. Siewert JR, Bottcher K, Stein HJ, et al: Relevant prognostic factors in gastric cancer: Ten-year results of the German Gastric Cancer Study. Ann Surg 228:449-461, 1998

28. Hyung WJ, Noh SH, Yoo CH, et al: Prognostic significance of metastatic lymph node ratio in T3 gastric cancer. World J Surg 26:323-329, 2002

29. Berger AC, Watson JC, Ross EA, et al: The metastatic/examined lymph node ratio is an important prognostic factor after pancreaticoduodenectomy for pancreatic adenocarcinoma. Am Surg 70: 235-240, 2004

30. Herr HW: Superiority of ratio based lymph node staging for bladder cancer. J Urol 169:943-945, 2003

31. Tsuchiya A, Kanno M, Abe R: The impact of lymph node metastases on the survival of breast cancer patients with ten or more positive lymph nodes. Surg Today 27:902-906, 1997

32. Walker MJ, Osborne MD, Young DC, et al: The natural history of breast cancer with more than 10 positive nodes. Am J Surg 169:575-579, 1995

33. Overgaard M, Hansen PS, Overgaard J, et al: Postoperative radiotherapy in high-risk premenopausal women with breast cancer who receive adjuvant chemotherapy: Danish Breast Cancer Cooperative Group 82b Trial. N Engl J Med 337:949-955, 1997

34. Grills IS, Kestin LL, Goldstein N, et al: Risk factors for regional nodal failure after breast-conserving therapy: Regional nodal irradiation reduces rate of axillary failure in patients with four or more positive lymph nodes. Int J Radiat Oncol Biol Phys 56:658-670, 2003

35. Knoop AS, Bentzen SM, Nielsen MM, et al: Value of epidermal growth factor receptor, HER2, p53, and steroid receptors in predicting the efficacy of tamoxifen in high-risk postmenopausal breast cancer patients. J Clin Oncol 19:3376-3384, 2001

36. Gürkaynak M, Yildiz F, Atahan IL: T3N0M0 Breast cancer patients: A subgroup with favorable prognosis. Turk J Cancer 33:144-149, 2003 http://www.turkjcancer.org/pdf/pdf_TJC_346.pdf

37. Germain I, Fortin A, Dagnault A, et al: The value of the number of nodes removed (or the ratio of involved nodes) as a prognostic factor in breast cancer. Proc Am Soc Clin Oncol 23:29, 2004 (abstr 611)

38. Holländer N, Schumacher M: On the problem of using 'optimal' cutpoints in the assessment of quantitative prognostic factors. Onkologie 24:194-199, 2001

39. Good PI: Resampling Methods: A Practical Guide to Data Analysis (ed 3). Boston, MA, Birkhauser, 2006

40. Guern AS, Vinh-Hung V: [Statistical distribution of involved axillary lymph nodes in breast cancer]. Bull Cancer 95:449-455, 2008

41. Fedorov V, Mannino F, Zhang R: Consequences of dichotomization. Pharm Stat [epub ahead of print on April 3, 2008]

42. Harrell FE Jr: Problems Caused by Categorizing Continuous Variables. Nashville, TN, Department of Biostatistics, Vanderbilt University, 2008. http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/CatContinuous

43. Austin PC, Brunner LJ: Inflation of the type I error rate when a continuous confounding variable is categorized in logistic regression analyses. Stat Med 23:1159-1178, 2004

44. Morgan TM, Elashoff RM: Effect of categorizing a continuous covariate on the comparison of survival time. J Am Stat Assoc 81:917-921, 1986

45. Schmoor C, Schumacher M: Effects of covariate omission and categorization when analysing randomized trials with the Cox model. Stat Med 16:225-237, 1997

46. Taylor JMG, Yu M: Bias and efficiency loss due to categorizing an explanatory variable. J Multivar Anal 83:248-263, 2002

47. Cserni G, Burzykowski T, Vinh-Hung V, et al: Axillary sentinel node and tumour-related factors associated with non-sentinel node involvement in breast cancer. Jpn J Clin Oncol 34:519-524, 2004

48. Farshid G, Pradhan M, Kollias J, et al: A decision aid for predicting non-sentinel node involvement in women with breast cancer and at least one positive sentinel node. Breast 13:494-501, 2004

49. Goyal A, Douglas-Jones A, Newcombe RG, et al: Predictors of non-sentinel lymph node metastasis in breast cancer patients. Eur J Cancer 40:1731-1737, 2004

50. Menes TS, Tartter PI, Mizrachi H, et al: Breast cancer patients with pN0(i+) and pN1(mi) sentinel nodes have high rate of nonsentinel node metastases. J Am Coll Surg 200:323-327, 2005

51. Barranger E, Coutant C, Flahault A, et al: An axilla scoring system to predict non-sentinel lymph node status in breast cancer patients with sentinel lymph node involvement. Breast Cancer Res Treat 91:113-119, 2005

52. Chagpar AB, Scoggins CR, Martin RC, et al: Prediction of sentinel lymph node-only disease in women with invasive breast cancer. Am J Surg 192:882-887, 2006

53. Wada N, Imoto S, Yamauchi C, et al: Predictors of tumour involvement in remaining axillary lymph nodes of breast cancer patients with positive sentinel lymph node. Eur J Surg Oncol 32:29-33, 2006

54. Dauphine CE, Haukoos JS, Vargas MP, et al: Evaluation of three scoring systems predicting non sentinel node metastasis in breast cancer patients with a positive sentinel node biopsy. Ann Surg Oncol 14:1014-1019, 2007

55. van Deurzen CH, van HR, Hobbelink MG, et al: Predictive value of tumor load in breast cancer sentinel lymph nodes for second echelon lymph node metastases. Cell Oncol 29:497-505, 2007

56. Pal A, Provenzano E, Duffy SW, et al: A model for predicting non-sentinel lymph node metastatic disease when the sentinel lymph node is positive. Br J Surg 95:302-309, 2008

57. Paik S, Tang G, Shak S, et al: Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 24:3726-3734, 2006

58. Braun S, Naume B: Circulating and disseminated tumor cells. J Clin Oncol 23:1623-1626, 2005

59. Pachmann K, Camara O, Kavallaris A, et al: Monitoring the response of circulating epithelial tumor cells to adjuvant chemotherapy in breast cancer allows detection of patients at risk of early relapse. J Clin Oncol 26:1208-1215, 2008

Acknowledgment

We thank Stina Blagojevic for editorial help and Phillip I. Good, PhD, for statistical advice.